

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: Apr11 22, 2002, 18:19:47 ; Search time 116.57 Seconds
(without alignments)
2391.648 Million cell updates/sec

Title: US-09-609-146-24

Perfect score: 1231
Sequence: 1 gttgtgatttaagtcag.....aggagtgtcagaagcctc 1231

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 11323899 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Issued_Patents_NA.*
2: /cgn2_6/ptodata/1/lna/5A.COMB.seq:*
3: /cgn2_6/ptodata/1/lna/5B.COMB.seq:*
4: /cgn2_6/ptodata/1/lna/6A.COMB.seq:*
5: /cgn2_6/ptodata/1/lna/6B.COMB.seq:*
6: /cgn2_6/ptodata/1/lna/Dockfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	133.2	10.8	1092	US-09-077-675A-15	Sequence 15, Appl
2	132.8	10.8	1088	US-09-077-675A-6	Sequence 6, Appl
3	128	10.4	1063	US-09-077-675A-1	Sequence 1, Appl
4	103.4	8.4	836	US-09-077-675A-11	Sequence 11, Appl
5	103.2	8.3	3129	US-09-077-675A-14	Sequence 14, Appl
6	100.2	8.1	1122	US-09-077-675A-9	Sequence 9, Appl
7	98.4	8.0	1029	US-09-077-675A-4	Sequence 4, Appl
8	95.6	7.8	1161	US-08-086-439C-2	Sequence 2, Appl
9	95.6	7.8	1367	US-08-434-877-2	Sequence 3, Appl
10	95.6	7.8	1370	US-08-475-742-3	Sequence 1, Appl
11	95.6	7.8	1370	US-08-056-051-1	Sequence 17, Appl
12	95.6	7.8	1370	US-07-928-611-17	Sequence 17, Appl
13	95.6	7.8	1370	US-08-487-811A-17	Sequence 17, Appl
14	95.6	7.8	1370	US-09-060-694-17	Sequence 17, Appl
15	95.6	7.8	1370	PCT-US93-07370-17	Sequence 17, Appl
16	95.6	7.8	1466	US-08-056-051-3	Sequence 3, Appl
17	95.6	7.8	1466	US-07-928-611-19	Sequence 19, Appl
18	95.6	7.8	1466	US-08-487-811A-19	Sequence 19, Appl
19	95.6	7.8	1466	US-09-060-694-19	Sequence 19, Appl
20	95.6	7.8	1466	PCT-US93-07370-19	Sequence 19, Appl
21	95.6	7.8	1610	US-08-056-051-5	Sequence 5, Appl
22	95.6	7.8	1610	US-07-928-611-21	Sequence 21, Appl
23	95.6	7.8	1610	US-08-487-811A-21	Sequence 21, Appl
24	95.6	7.8	1610	US-09-060-694-21	Sequence 21, Appl
25	95.6	7.8	1610	PCT-US93-07370-21	Sequence 21, Appl
26	91.6	7.4	1529	US-08-858-876A-3	Sequence 3, Appl
27	91.6	7.4	1529	US-09-472-880-3	Sequence 3, Appl

28	79.6	6.5	283	US-08-993-088A-4	Sequence 4, Appl
29	76.8	6.2	1205	US-08-417-103-13	Sequence 13, Appl
30	76.8	6.2	1634	US-07-816-283-1	Sequence 1, Appl
31	76.8	6.2	1634	US-08-417-103-1	Sequence 1, Appl
32	70.2	5.7	1261	US-08-475-742-16	Sequence 16, Appl
33	70	5.7	1410	US-08-147-592A-1	Sequence 13, Appl
34	69.4	5.6	936	US-08-288-663A-13	Sequence 2, Appl
35	69.4	5.6	1194	US-08-288-663A-2	Sequence 3, Appl
36	69.4	5.6	1228	US-08-858-876A-1	Sequence 1, Appl
37	68.8	5.6	1575	US-09-472-880-1	Sequence 1, Appl
38	68.8	5.6	1575	US-08-832-399-1	Sequence 1, Appl
39	68.6	5.6	1342	US-09-372-498-1	Sequence 1, Appl
40	68.6	5.5	1265	US-07-816-283-3	Sequence 3, Appl
41	67.6	5.5	1265	US-08-417-103-3	Sequence 11, Appl
42	67.6	5.5	1085	US-08-466-103A-11	Sequence 6, Appl
43	66.8	5.4	1085	US-08-993-088A-6	Sequence 7, Appl
44	66.2	5.4	1164	US-08-722-001-7	
45	66.2	5.4	1601	US-08-722-001-7	

ALIGNMENTS

RESULT 1
US-09-077-675A-15
; Sequence 15, Application US/09077675A
; Patent No. 6242199
GENERAL INFORMATION:
APPLICANT: Pail, Lee-Yuh
APPLICANT: Feigner, Scott C.
APPLICANT: Howard, Andrew D.
APPLICANT: Peng, Sheng-Shung
APPLICANT: Van Der Ploeg, Leonardus H.T.
TITLE OF INVENTION: RECEPTOR ASSAY
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: P. O. Box 2000, 126 E. Lincoln Ave.
CITY: Rahway
STATE: NJ
COUNTRY: USA
ZIP: 07065-0900
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/077, 675A
FILING DATE: 3-JUN-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Cocuzzo, Anna L.
REGISTRATION NUMBER: 42,452
REFERENCE/DOCKET NUMBER: 19590P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 732-594-1273
TELEFAX: 732-594-4720
TELEX:
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 1092 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-077-675A-15
Query Match 10.8%; Score 133.2; DB 4; Length 1092;

RESULT 4
US-09-077-675A-11
Sequence 11, Application US/09077675A
Patent No. 6242199
GENERAL INFORMATION:
APPLICANT: Pal, Lee-Yuh
APPLICANT: Feighner, Scott C.
APPLICANT: Howard, Andrew D.
APPLICANT: Pong, Sheng-Shung
APPLICANT: Van Der Ploeg, Leonardus H.T.
TITLE OF INVENTION: RECEPTOR ASSAY
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000, 126 E. Lincoln Ave.
CITY: Rahway
STATE: NJ
COUNTRY: USA
ZIP: 07065-0900
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/077,675A
FILING DATE: 3-JUN-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Cocuzzo, Anna L.
REGISTRATION NUMBER: 42,452
REFERENCE/DOCKET NUMBER: 19590P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 732-594-1273
TELEFAX: 732-594-4720
TELEX:
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 836 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-077-675A-11

Query Match 8.4%; Score 103.4; DB 4; Length 836;
Best Local Similarity 49.1%; Pred. No. 1.1e-19;
Matches 362; Conservative 0; Mismatches 366; Indels 9; Gaps 3;
QY 275 ctggtctgctcttggtggtatgctcttggaatctcaagatgtgacaaatcccttc 334
DB 3 ctctctatcttctctgacacccctgacctgcttgcctctgagagatccggccttg 62
QY 335 ctgtctgggagctgtggaatctactcaagaagccctcttgagactgtgtcttgc 394
DB 63 AACTGGGGACCTCTCTGCAACTCTTCCAAATTCGTAGTAGAGAGCTGCACCTGAGCC 122
QY 395 tccattctcagtgtagcaagcttagagcgtatgtggcaattgtccacccttc 454
DB 123 ACGGTCTCTACATCAACAGGCTGAGGCTGAGCGCTACTTCCCATCTCTCCACTC 182
QY 455 cgaagcaagcttgagagacagcgagcgagccctcaagatccctcagctagcttgagc 514
DB 183 CGGGCAAGAGGTGGTGACCAAGGGGGGAGCTGTGATCTTCGTATCTGGGGC 242
QY 515 ttctctgtgtcttcttcttgcacataaccagatcatgtgcatcaagttccagcatt 574
DB 243 GTGGCCTTGTGACGCGGGGCCATCTTGTGCTAGTCGGGGTGAGGACGAGAACGGC 302

QY 575 cccaaggggtctcctgtaactggtctcaagccactgcaagctcaacaaacccatgtgggtg 634
DB 303 ACCGACCCCTTGGGACACCAAGAGTGGCCGCCACCGAGTTTGGCGCTTGAGACTG 362
QY 635 tataacttgatcaaccaagctaccagcttctctcttctatccctccaaatgacccctc 694
DB 363 CTCACGGTATGAGTGGTGGGTCACAGATCTTCTCTCTCTCTCTCTCTCTCTCTCT 419
QY 695 agcgtctctactactcaactcaatggggtctcaggtctgaagagatgaatcccttgagcgaac 754
DB 420 ACGGTCTCTTACAGTCTCAACGAGAGAACCTGTGGCGGAGAGCGCGGCGGATGCTGTC 479
QY 755 aaagtgcgtgataatcaacagaccctcagaagaatcaatcaacaaagtgtgtgtc 814
DB 480 ---GTGGGTCTTCGCTCAAGGACAGAACCAACAGAACCTGAAATGCTGCTGTA 536
QY 815 ttggtctcgtgttgccatctgctgagacccctctccatgtgacg---gctctcttc 871
DB 537 GTGGTGTGCTCTTCATTCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 596
QY 872 agcttgtggaagatggaagatgctccctgtgctgtgttcaactcatctatgtgta 931
DB 597 AAATCTTGAAGCTGCTGCTCTTGAAGATTGCTCAGATCAGCACTGCAACTGCACTGTC 656
QY 932 tcaagtgctctcttctatctgaagctcgcggtcaacccatctataacccctctgtc 991
DB 657 TCCTTGTCTCTCTTCTTACCTCACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 716
QY 992 cggcgcttccggcgagc 1008
DB 717 AAGAAGTACCGGGTGGC 733

RESULT 5
US-09-077-675A-14
Sequence 14, Application US/09077675A
Patent No. 6242199
GENERAL INFORMATION:
APPLICANT: Pal, Lee-Yuh
APPLICANT: Feighner, Scott C.
APPLICANT: Howard, Andrew D.
APPLICANT: Pong, Sheng-Shung
APPLICANT: Van Der Ploeg, Leonardus H.T.
TITLE OF INVENTION: RECEPTOR ASSAY
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000, 126 E. Lincoln Ave.
CITY: Rahway
STATE: NJ
COUNTRY: USA
ZIP: 07065-0900
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/077,675A
FILING DATE: 3-JUN-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Cocuzzo, Anna L.
REGISTRATION NUMBER: 42,452
REFERENCE/DOCKET NUMBER: 19590P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 732-594-1273
TELEFAX: 732-594-4720
TELEX:
INFORMATION FOR SEQ ID NO: 14:

Best Local Similarity 50.7%; Pred. No. 1.9e-17;
Matches 230; Conservative 0; Mismatches 224; Indels 0; Gaps 0;

```

Oy 138 cggctcgtggtccatgcgtcgtatctcctcgtggtggtatggtgcaatctctcgtggt 197
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 104 cggccctggtggggggcgtcgtcgtatcgtggtggtggtggtggtggtggtggtggt 163
Oy 138 gcatggtggttgcgcagatctcgtatgaaagacaccccaactactatctctcagct 257
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 164 gcgtagcgtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 223
Oy 228 tggcagctcagatctggtcgtcgtcgtcgtggtggtggtggtggtggtggtggtggt 317
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 224 tggcggcggcggcggcggcggcggcggcggcggcggcggcggcggcggcggcggcgg 283
Oy 318 ggcacaattaccccttctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 377
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 284 tccagcgtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 343
Oy 378 agactggtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 437
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 344 tcatcgtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 403
Oy 438 ccatgtcagcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 497
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 404 cggcgtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 463
Oy 488 tcaagctggtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 557
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 464 tggcggcggcggcggcggcggcggcggcggcggcggcggcggcggcggcggcggcgg 523
Oy 558 tcaagctggtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 591
Db 524 acgtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 557

```

RESULT 9

US-08-434-877-2

Sequence 2, Application US/08434877

Patent No. 5721132

GENERAL INFORMATION:

APPLICANT: Chio, Christopher L.

APPLICANT: Huff, Rita M.

TITLE OF INVENTION: A Synthetic Gene for D4 Dopamine

TITLE OF INVENTION: Receptors

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESSES:

ADDRESSEE: The Upjohn Company, Corp. Intellectual

ADDRESSEE: Property Law

STREET: 301 Henrietta Street

CITY: Kalamazoo

STATE: Michigan

COUNTRY: USA

ZIP: 49001

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette (DS,HD)

COMPUTER: Gateway 2000 P5-90

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/434,877

FILING DATE: 1 July 1993

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Darnley Jr., James D.

REGISTRATION NUMBER: 33,673

REFERENCE/DOCKET NUMBER: 4700 DVI

TELECOMMUNICATION INFORMATION:

TELEPHONE: 616-385-5210

TELEFAX: 616-385-6897

TELEX: 224401

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1161 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-434-877-2

Query Match 7.8%; Score 95.6; DB 1: Length 1161;
Best Local Similarity 50.7%; Pred. No. 1.9e-17;
Matches 230; Conservative 0; Mismatches 224; Indels 0; Gaps 0;

```

Oy 138 cggctcgtggtccatgcgtcgtatctcctcgtggtggtatggtgcaatctctcgtggt 197
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 104 cggccctggtggggggcgtcgtcgtatcgtggtggtggtggtggtggtggtggtggt 163
Oy 138 gcatggtggttgcgcagatctcgtatgaaagacaccccaactactatctctcagct 257
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 164 gcgtagcgtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 223
Oy 228 tggcagctcagatctggtcgtcgtcgtcgtggtggtggtggtggtggtggtggtggt 317
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 224 tggcggcggcggcggcggcggcggcggcggcggcggcggcggcggcggcggcggcgg 283
Oy 318 ggcacaattaccccttctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 377
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 284 tccagcgtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 343
Oy 378 agactggtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 437
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 344 tcatcgtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 403
Oy 438 ccatgtcagcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 497
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 404 cggcgtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 463
Oy 488 tcaagctggtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 557
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 464 tggcggcggcggcggcggcggcggcggcggcggcggcggcggcggcggcggcggcgg 523
Oy 558 tcaagctggtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 591
Db 524 acgtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 557

```

RESULT 10

US-08-475-742-3

Sequence 3, Application US/08475742

Patent No. 6121015

GENERAL INFORMATION:

APPLICANT: O'Malley, Karen L.

APPLICANT: Todd, Richard D.

TITLE OF INVENTION: Gene Encoding the Rat Dopamine D4 Receptor

FILE REFERENCE: WU 102 CON DIV

CURRENT APPLICATION NUMBER: US/08/475,742

EARLIER FILING DATE: 1995-06-07

EARLIER APPLICATION NUMBER: US 08/261,293

EARLIER FILING DATE: 1994-06-16

EARLIER APPLICATION NUMBER: US 08/014,013

NUMBER OF SEQ ID NOS: 16

SOFTWARE: Patent Ver. 2.0

SEQ ID NO 3

LENGTH: 1367

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc.feature

LOCATION: (1)-(1367)

OTHER INFORMATION: D4 Dopamine Receptor cDNA

PUBLICATION INFORMATION:

AUTHORS: Van Tol, H. H.

AUTHORS: Bunzow, J. R.

: TITLE: Cloning of the gene for a human dopamine D4 receptor
 : TITLE: with high affinity for the antipsychotic clozapine
 : JOURNAL: Nature
 : VOLUME: 350
 : PAGES: 610-614
 : DATE: 1991
 : JS-08-475-742-3

Query Match	7.8%	Score 95.6;	DB 3;	Length 1367;
Best Local Similarity	50.7%	Pred. No. 2.1e-17;		
Matches 230; Conservative	0;	Mismatches 224;	Indels 0;	Gaps 0

Oy	138	cggctctctgagccatctgagatctcttcacgctgctgggggagatagagcaatcttcgtgt	197
Db	210	cggcgctctgtaggggggctgtctctcatctcggctggctggctggagatctgctcgtgt	268
Oy	198	gcatctgtatctgtccggaatcagaactcttgaaagacaccacaaatactatctcttcagct	257
Db	270	ggcttgagagctgtgacacacgagctggccctctgacacgcccacaaactctcttcactgtggcc	329
Oy	258	tggcagctctcagatcctctgctgtgtcctgtctctctggggatagctctctggaatctacgagatgt	317
Db	330	tggcgccgcgcgcgaactctctctctctcctcctcctctgctgtgcgcgcctctctctgtactaccgagtg	389
Oy	318	ggcaacaaatcaacccctctctctctctggggcccttgtagatctacttcaagacagccctcttcg	377
Db	390	tcccgagctgagctgctgtctgtctctgagcccccgcctctgtcgaagccctcatctgacatgagacg	449
Oy	378	agactctgtctctctctctcacaatctctcagctgtcatcccaagctctatgacgctacatgtgtg	437
Db	450	tcatctctgttgacacgcgcctcactctctcaaacctctgtcgcacatcagctgtgacagcttcgt	509
Oy	438	ccatctgtccaaacccctctccgagacacagctctggagagacagccggccgagccctcaagatcc	497
Db	510	ccgtgtgcgcgtgcgcgtgcgtacacacccgacggctctggagagccggtccgagctctgtctca	568
Oy	498	tcaagctctatctctggagactctctctctgtgtctctctctctctctctctcaatcacaatccatctgaca	557
Db	570	tctggcgcaacgtgtgctgtctgtctcgcggctggctggctggccgcttactctgtgacctcaacg	629
Oy	558	tcaagctctcagaactctctccaaagctctccgt	591
Db	630	acgtctgcgcgcgcgcgaacccctcgtctgtctgcgcct	663

RESULT 11
US-08-056-051-1

```

1  GENERAL INFORMATION:
2  APPLICANT: Grandy, David K
3  APPLICANT: Bunzow, James R
4  APPLICANT: Civellet, Olivier
5  APPLICANT: Van Tol, Hubert H.-M.
6  TITLE OF INVENTION: A No. 551663e1 Human Dopamine Receptor and Uses
7  NUMBER OF SEQUENCES: 6
8  CORRESPONDENCE ADDRESS:
9

```

```

1  COMPUTER READABLE FORM:
2  MEDIUM TYPE: Floppy disk
3  COMPUTER: IBM PC compatible
4  OPERATING SYSTEM: PC-DOS/MS-DOS
5  SOFTWARE: PatentIn Release #1.0, Version #1.2
6  CURRENT APPLICATION DATA:
7  APPLICATION NUMBER: US/08/056,051
8  FILING DATE: 19930429
9  CLASSIFICATION: 435

```

```

1  ATTORNEY/AGENT INFORMATION:
2  NAME: NO. 5516683nan, Kevin E
3  REGISTRATION NUMBER: 35,303
4  REFERENCE/DOCKET NUMBER: 90,1092-C
5  TELECOMMUNICATION INFORMATION:
6  TELEPHONE: 312-715-1000
7  TELEFAX: 312-715-1234
8  TELER: 910-221-5317
9  INFORMATION FOR SEQ ID NO: 1:
10  SEQUENCE CHARACTERISTICS:
11  LENGTH: 1370 base pairs
12  TYPE: NUCLEIC ACID
13  STRANDEDNESS: single
14  TOPOLOGY: linear
15  MOLECULE TYPE: cDNA
16  FEATURE:
17  NAME/KEY: 5'UTR
18  LOCATION: 1..103
19  FEATURE:
20  NAME/KEY: 3'UTR
21  LOCATION: 1268..1370
22  FEATURE:
23  NAME/KEY: CDS
24  LOCATION: 104..1267
25  US-08-056-051-1

```

Query Match	7.8%;	Score 95.6;	DB 1;	Length 1370;
Best Local Similarity	50.7%;	Pred. NO. 2.1e-17;		
Matches 230; Conservative	0;	Mismatches 224;	Indels 0;	Gaps 0;

Oy	138	cggtctctgtgactatbpcgtcttgccttccctcgtggtgggtgaatgagcaatcttcgtgt	197
Db	210	CGGCCCTGTGTGGGGGGCTGCTGCTCATTCGGCGGGTCTCGGGGAACTCCGCTGCTG	269
Oy	198	gcatgtgtgatgtctcgcgaatcagaacttltgaagacacccaccatactatcttcgaact	257
Db	270	GGGTAGAGGTGGCCACCGACGAGCGGCCCTGCGAGAGCCCAACTCTCTCATGTGTGACC	329
Oy	258	tggcaagctcaagactctgtgttctgtctcttcttgggagagactcttggaaatcttaagagatgt	317
Db	330	TGGCGCGCGGACACTCTCTCTGCTCTCTCCGTGGTGGCTGCGGCTTCTGTACTCTCCGAGG	389
Oy	318	ggcaaatctaccccttccctctgtcttgcggccttbgatgtactacttaacagacagccctctcg	377
Db	390	TCCAGGGTGGGCGGAGGTGCTGTAGGCCCGCGCTGTGCGAGCGCCCTCATGGCCATGGAGG	449
Oy	378	agaactgtgtcttgccttgccttccatcttcagtygtcacacagtgtagcgtatagagcgtatgtg	437
Db	450	TCATGCTGTGACCGCCCTCCATCTTCAACAGTGCAGCCATTCAGGTGGAGAGGTTGCTGG	509
Oy	438	gcatgtctcaccccttccgcagacgaagcttggagagacagcgcgagcagggccctcaagatcc	497
Db	510	CCGTGGCCGTGCTGCTGCGCTACAAACGGGCAAGGTGGAGACCCCGCAAGCTGCTGCTCA	569
Oy	498	tcaagcttagtcttggagcttctctgtgtgttcttcttccccaataccagatcatcattggca	557
Db	570	TGCGGGCCACAGCTGGCTGCTGTCCCGCGGAGGTGGCGGCGCCGTAAGTGTGGCGCTCAACG	629
Oy	558	tcaagcttcagcaacttcccaaggggtccctcogt	591
Db	630	ACGTGCGCGCGGACCGCCCGCTGTGGCGCCT	663

```

RESULT 12
US-07-928-611-17
: Sequence 17, Application US/07928611
: Patent No. 5569601
:
: GENERAL INFORMATION:
:
: APPLICANT: Van Tol, Hubert H.M.
: APPLICANT: Civeill, Olivier
: TITLE OF INVENTION: A No. 5569601el Human Dopamine Receptor and Uses
: NUMBER OF SEQUENCES: 22

```


Best Local Similarity 50.7%; Pred. No. 2,1e-17;
Matches 230; Conservative 0; Mismatches 224; Indels 0; Gaps 0;

```
OY 138 cggctgctgtgacctgacctgacctgtaccttctctgtggtgtaatggcaatctctgtgt 197
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 210 CGGCGGTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 198 gcaatgtagtctgcgacatcagacttgaagacacacacacacacacacacacacacac
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 270 GCGTGAGCGTGCGCACCGGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 258 tggcagctcagatctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 330 TGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 318 ggcacaaataacaccttccctgctcggtcggtcggtcggtcggtcggtcggtcggtcggt
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 390 TCCAGGCGTGGCGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCG
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 378 agactgtgtcttgccttgccttgccttgccttgccttgccttgccttgccttgccttgc
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 450 TCATGCTGTGACCGCGCTTCATCTTAACCTGTGCGCCCATCAAGCTGACAGGTTGCTGG
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 438 ccattgtccacccttcccttcccttcccttcccttcccttcccttcccttcccttcccttcc
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 510 CCGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 498 tcaagctagctggaagctctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 570 TCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 558 tcaagttccagcacttcccaaggggtcctcgt 591
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 630 ACGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: April 22, 2002, 18:20:03
Job time: 9179 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2002, 14:17:43 ; Search time 25.8 Seconds
(without alignments)
1166.238 Million cell updates/sec

Title: US-09-609-146-25
Perfect score: 2076
Sequence: 1 MGKLENSWTHDPLMKYLYNS.....QSSIHNTNLTPACAGEVP 395

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Prctd. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	513	24.7	424	2 JH0164	neurotensin recept
2	499	24.0	418	2 S29506	neurotensin recept
3	457.5	22.0	378	2 T15816	hypothetical prote
4	430.5	20.7	416	2 S68822	neurotensin recept
5	416	20.0	398	2 JN0708	thyrotropin-releas
6	412	19.8	393	2 A39251	thyrotropin-releas
7	409	19.7	402	2 I56595	neurokinin 2 recep
8	403.5	19.4	411	2 I56444	thyrotropin-relea
9	403.5	19.4	412	2 S23436	thyrotropin recep
10	402.5	19.3	392	2 S65693	opioid receptor mu
11	400.5	19.3	400	2 I56553	mu opiate receptor
12	399.5	19.2	384	2 S00516	neurokinin 2 recep
13	394	19.0	398	1 J01059	neurokinin 3 recep
14	387.5	18.7	385	2 S55524	neurokinin 3 recep
15	387	18.6	398	2 A57510	mu opioid receptor
16	381.5	18.4	394	2 J07209	galanin receptor -
17	380.5	18.3	452	2 A34916	neurokinin 3 recep
18	380	18.3	398	2 I56517	mu opioid receptor
19	373	18.0	384	2 S20303	neurokinin 2 recep
20	372	17.9	398	2 I56504	mu opioid receptor
21	368	17.7	390	2 A36737	neurokinin 2 recep
22	366.5	17.7	465	1 J01517	neurokinin 3 recep
23	363.5	17.5	384	2 A47249	brain-specific som
24	362.5	17.5	363	2 I57940	somatosstatin recep
25	362	17.4	380	2 S36143	kappa opioid recep
26	362	17.4	384	2 I57957	neurokinin 2 recep
27	360	17.3	380	2 A35259	kappa opioid recep
28	359.5	17.3	380	2 JC2338	kappa opioid recep
29	359	17.3	369	2 B41795	somatosstatin recep

30	359	17.3	440	2 A44081	kappa-type opioid
31	357	17.2	372	2 S34592	delta opioid recep
32	357	17.2	380	2 A48227	kappa opioid recep
33	357	17.2	504	2 A41783	tachykinin recepto
34	356	17.1	388	2 JN0605	somatosstatin recep
35	356	17.1	399	2 S29480	bombesin receptor
36	355	17.1	369	2 J02083	somatosstatin recep
37	354.5	17.1	352	2 J02096	thyrotropin releas
38	354.5	17.1	399	2 A46632	bombesin-like pept
39	353.5	17.0	384	2 JC4629	somatosstatin recep
40	351.5	16.9	372	2 I38532	delta opioid recep
41	351	16.9	369	2 A45291	somatosstatin recep
42	350	16.9	369	2 D41795	somatosstatin recep
43	349	16.8	380	2 JC2434	kappa opioid recep
44	347.5	16.7	418	2 A46326	somatosstatin recep
45	347	16.7	346	2 S29248	somatosstatin recep

ALIGNMENTS

RESULT 1
JH0164
neurotensin receptor - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #extl_change 17-Mar-2000
C:Accession: JH0164
R:Tanaka, K.; Masu, M.; Nakanishi, S.
Neuron 4, 847-854, 1990
A>Title: Structure and functional expression of the cloned rat neurotensin receptor.
A:Reference number: JH0164; MUID:90297956
A:Accession: JH0164
A:Molecule type: mRNA
A:Residues: 1-424 <TAN>
C:Comment: Neurotensin receptor belongs to the family of G protein-coupled receptor.
ter (neuromodulator in the brain and as a hormone) cellular mediator in peripheral t
C:Superfamily: vertebrate rhodopsin
C:Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein
F:65-87/Domain: transmembrane #status predicted <TM1>
F:97-121/Domain: transmembrane #status predicted <TM2>
F:144-165/Domain: transmembrane #status predicted <TM3>
F:189-210/Domain: transmembrane #status predicted <TM4>
F:236-260/Domain: transmembrane #status predicted <TM5>
F:309-330/Domain: transmembrane #status predicted <TM6>
F:348-372/Domain: transmembrane #status predicted <TM7>
F:4,38,42/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 24.7%: Score 513; DB 2; Length 424;
Best local similarity 31.6%: Pred. No. 1,1e-35;
Matches 119; Conservative 76; Mismatches 113; Indels 68; Gaps 13;
QY 19 NSTEEYLHLGPRKSLDP-----VSVAVALIFLVGNGNLLVCNIVRH--QT 67
DB 42 NTSESDPA--GP-NSDLVDNTDIYSKVLATYALFVVGTVGNSVTAFTRKSLQS 97
QY 68 LKPTNYLFLSVSDLVLLGMPLEIYE-MHMYNPELEGPGCC---YKRTALFEYVFC 123
DB 98 LOSTVHYHLGSLSDLLILLAMPVELYNFIWHPHMGAGDGCRCGYF--LRDCTY 154
QY 124 ASLISVTVVERVVAIVHPRAKLESTRRALRIISLWSFSVFSLPMTSHIGIKFQH 183
DB 155 APLNANVASVERLYLACHCFKAKTLMRSRRTKFKFISAILASALLAIPMLFTMGL--QN 212
QY 184 FPNSSVPGSATCTYTKPMVNVNLIIOANSFLFYILPMTLISVLYLMGLRKDESLA 243
DB 213 RSGDGTGPGGLVCTPIVDATVVKVIOVNTFMSFLFPMIVISIL-----NTVIA 261
QY 244 NKVAVNHR-----PBRKVTK-----MLFVLVVAICW 273
DB 262 NKLIVNVAHQAEQRCVCTVNGHLESHSTFNMTTEPQVALRHGVLAIVAVIAFVWC 321
QY 274 TPRHVDRLFFSFV--EEWTESLAAVFNLIVHVGVPFYLSAVANPIIYNLSRRFRAFR 331

Db 322 LPVHVRIMCYISDEQWTFLEDFEYHYFMALNFYSSAINPILYNLSANFRQVFL 381
 QY 332 NVVSPYCK-WCHPRHR 346
 Db 382 STLACLCPGWHRHRKK 397

RESULT 2
 S29506
 Query Match 24.0%; Score 499; DB 2; Length 418;
 Best Local Similarity 33.0%; Pred. No. 1,6e-34;
 Matches 116; Conservative 65; Mismatches 103; Indels 68; Gaps 12;

neurotensin receptor - human
 C:Species: Homo sapiens (man)
 C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 17-Mar-2000
 C:Accession: S29506
 R:Vita, N.; Laurent, P.; Lefort, S.; Chalou, P.; Dumont, X.; Kaghad, M.; Gully, D.; Le F
 FERS Lett. 317, 139-142, 1993
 A:Title: Cloning and expression of a complementary DNA encoding a high affinity human ne
 A:Reference number: S29506; MUID:93154505
 A:Accession: S29506
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-418 <VIT>
 A:Cross-references: EMBL:X70070; NID:g35020; PIDN:CAA49675.1; PID:g35021
 C:Superfamily: vertebrate rhodopsin
 C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 24.0%; Score 499; DB 2; Length 418;
 Best Local Similarity 33.0%; Pred. No. 1,6e-34;
 Matches 116; Conservative 65; Mismatches 103; Indels 68; Gaps 12;

QY 40 VSAVALIFLVGVNGLVCMVIVRH---OTLKTPTNYLFSLAVSDLVLLGMPLEIY 96
 Db 66 VFAVYALFVGVGVNMTVFTLARKKSLQSLQSTVHYHGLSLDILLLLAMPVELY 125
 QY 97 E-WMHNPFLFGVGC---YFKTALFETVCFASLSTVTVSERYVAIVHPRAKLESTR 152
 Db 126 NFTWVHHPMAFGDAGCGYF---LRDACTYATALNVASLSVERYLAIQHPRFAKTLMSR 182
 QY 153 RRALRILSLVMSFSVSESLP-----NLSHGIRKQHPNPGSSVGSATCVTKPMV 204
 Db 183 SRKKRFLSAIMLASALLTVMLTMEQNSADG---QH-----AGGLVCTPTHTAT 232
 QY 205 YNLIGATSEFLYLPMTLISLVLYLGLRKDESLKANKAVNINR-----252
 Db 233 VKVVIQVNTFMSRIFPMVIVSL-----NTIIANKLVYMQAEEQGVCTVGC 281
 QY 253 -----PSR-----KSTKMLFVLVLPALCMTFPHVDRLFESFV--EEMTESLAAY 296
 Db 282 EHSTFSMALEPGRVQALRHGVRVLRRAVIAFVVCWLPYHRRRLMFCYISDEQWTFLEDF 341
 QY 297 FNLIHVSGVFYLLSSAVNPDIINLSRRPRAFRNVSPYCK-WCHPRHRP 347
 Db 342 YHFEYVNTALFYVSSITNIPILYNLSANFRHIFLATLACLCFVWRRRRKKRP 393

RESULT 3
 T15816
 hypothetical protein C48C5.1 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 17-Mar-2000
 C:Accession: T15816
 R:Pavellio, A.
 submitted to the EMBL Data Library, November 1995
 A:Description: The sequence of C. elegans cosmid C48C5.
 A:Reference number: Z18410
 A:Accession: T15816
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-378 <FAV>
 A:Cross-references: EMBL:U39994; NID:g1055102; PID:g1055105; PIDN:AA837017.1; GSPDB:GN00
 A:Experimental source: strain Bristol N2; clone C48C5
 C:Genetics:

A:Gene: CESP:C48C5.1
 A:Map position: X
 A:Introns: 53/3; 87/2; 136/2; 169/3; 209/3; 231/3; 259/3; 286/1; 327/1
 C:Superfamily: adenostine receptor A1

Query Match 22.0%; Score 457.5; DB 2; Length 378;
 Best Local Similarity 29.9%; Pred. No. 4,5e-31;
 Matches 109; Conservative 74; Mismatches 140; Indels 41; Gaps 10;

QY 14 LMKYLNSTEE-----YLAHLCGPKRS--DLSL-PYSVAVALI 47
 Db 2 LQACLNTTEQCCCLAFNCPYVSHSESEKAEYMEHCIFSKRALDVTLYKYVALYIFI 61
 QY 48 FIVGVGNLLVCMVIVHQTPTNYLFSLAVSDLVLLGMPLEIYEMHNPFLFG 107
 Db 62 FLVGVGNNTTTCVLMKHPMKTHASWYLNLAIVSDLVTLVCGIPFVMMNMNOYPPFP 121
 QY 108 PVGCYETALFETVCFASLSTVTVSERYVAIVHP-FRAKLESTRRALRILSLVMSFS 166
 Db 122 DYICNLKALAEETSSYSILILIFALERVAVACHPLFMKQOPKRNICTIIGFTWIFS 181
 QY 167 VVFSLEPNTSHGKE--QHFP---NGSSVPGSATCTVT-----KPMVYLLIQATSEFL 216
 Db 182 ILCAPEAIHHRADYIKMSWPGTDNRIPVKSXKCMIAVNEPDKLSTFKILFHFSAIAP 241
 QY 217 YILPMTLSVLYLMLGLRKDESLKANKAVANIHRSRSKSVTKMLFVLVLPALCMTF 276
 Db 242 FALPFTVILVIRACKVSSNRITQPE--LDITTELOWRINAIICAIYSAFFICLUP 299
 QY 277 HYDRLEFSEVEEMTESLAAYFNLIHVSGVFYLLSSAVNPDIINLSRRPRAFRNVSP 336
 Db 300 QIDRLLEFFEDN-EVILTWVNOQWYFISGFLPLATIIINIAVNLASSRRRAFKDILID 358
 QY 337 TCKW 340
 Db 359 YC-W 361

RESULT 4
 S68822
 neurotensin receptor 2, leucobastine-sensitive - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 04-Dec-1997 #sequence_revision 12-Dec-1997 #text_change 20-Jun-2000
 C:Accession: S68822
 R:Chalon, P.; Vita, N.; Kaghad, M.; Gulleriot, M.; Bonnin, J.; Delpech, B.; Le Fur, G
 FERS Lett. 366, 91-94, 1996
 A:Title: Molecular cloning of a leucobastine-sensitive neurotensin binding site.
 A:Reference number: S68822; MUID:96228041
 A:Accession: S68822
 A:Molecule type: mRNA
 A:Residues: 1-416 <CHA>
 A:Cross-references: GB:X97121; NID:g1483579; PIDN:CAA65787.1; PID:g1483580
 A:Experimental source: hypothalamus
 C:Superfamily: vertebrate rhodopsin
 C:Keywords: G protein-coupled receptor; transmembrane protein
 F:33-58/Domain: transmembrane #status predicted <TM1>
 F:70-91/Domain: transmembrane #status predicted <TM2>
 F:110-131/Domain: transmembrane #status predicted <TM3>
 F:155-175/Domain: transmembrane #status predicted <TM4>
 F:204-230/Domain: transmembrane #status predicted <TM5>
 F:296-315/Domain: transmembrane #status predicted <TM6>
 F:335-361/Domain: transmembrane #status predicted <TM7>

Query Match 20.7%; Score 430.5; DB 2; Length 416;
 Best Local Similarity 31.5%; Pred. No. 9,4e-29;
 Matches 113; Conservative 61; Mismatches 126; Indels 59; Gaps 11;

QY 44 YALIFLVGVNGLVCMVIVRHQTAKT-PINNYLFSLAVSDLVLLGMPLEIYE-WMHN 101
 Db 39 YSLIFAGTGNALSYHVILKARAGRGRLRYHVLALSALLLVSMPELNYNFWVSH 98

Db 135 FLCTFSRAKKIIFVMAFTSIYCMLEFLDLNISTYKNAVV-----SCGKISIRNYSP 190
 Qy 208 IIQATSLFYILPMTLSVLYL-----GLRLKRDSELEANKAVNIHR 252
 Db 191 IYLMDFCVFVVPMTLVGLGFARLFLNPIPSDEKSKMKKNOSIONK-NLNLNA 249
 Qy 253 -----PSRKSVTKMLFVLVFAICWTPFH---VDRLFES--EVEWTESLAAVFN 298
 Db 250 TNCNFTSVSSRKQVTKMLAVVILFALMMFRTLVVNSFLSSPQENK----- 300
 Qy 299 LHVSGVFFYLSAVNPITYNLSRRFRAFRNVSPCTCKCHPRHRPPAOKIT-- 356
 Db 301 -FLFCRICIYLNLSAIPVYNLMSQKFAFR-----KLCNCKOKPTEKAMYSVAL 352
 Qy 357 -----FLTECHLVELTE 368
 Db 353 NYSYIKESDRFSTLELDTIVTD 374

RESULT 7
 156595
 neurokinin 2 receptor - guinea pig
 C/Species: Cavia porcellus (guinea pig)
 C/Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 20-Apr-2000
 C/Accession: I56595
 R/Author: D.; Little, J.; Thomas, C.; Powell, S.; Downey-Jones, M.; Graham, A.
 J. Recept. Res. 14, 399-421, 1994
 A/Title: Isolation and characterization of neurokinin A receptor cDNAs from guinea-pig 1
 A/Reference number: I56595; MUID:95162423
 A/Accession: I56595
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-402 <RES>
 A/Cross-references: GB:S76553; NID:g913274; PIDN:AAB33553.1; PID:g913275
 C/Superfamily: neurokinin 1 receptor

Query Match 19.7%; Score 409; DB 2; Length 402;
 Best Local Similarity 29.3%; Pred. No. 5.9e-27;
 Matches 105; Conservative 65; Mismatches 144; Indels 44; Gaps 11;

Qy 41 SVATALLFLVGMGNLVCNIVYRHQTKPTNYTFLSLAVSDLLVLLGMPLEIYEMMH 100
 Db 37 ATAVLALVLAVTGNATVTITLHQRMRTVNYFIYNLALADLMAAFNAFNFVYASH 96
 Qy 101 NYPLFPGVCGYFRTALFETVCFASILSVTVSERYVAIVHPRAKLESTRRRALRILS 160
 Db 97 NT-WYGFARFCYFONLFPTITAMFVSYISMTALAIIDRYMALVHPQRLSAFSTKA--VIG 153
 Qy 161 LVMSFVVFSLPNTSINGIKFQHPNGSSVPSGATCTVT-----KPMVYNLIQATS 213
 Db 154 GIMVVALALAPQCFYSTI-----TDEGATKCVVAMPEDSRDSSLILHLV---I 202
 Qy 214 FLFYILPMTLSVLYLMGLKRLKRDSELEANKAVNI-HRPSKSVTKMLFVLVLAIC 272
 Db 203 VLYILPMTLVFVYVSIIGITLMRAVRPHQAHGANLRLHQAQKFKVTWLVVVTPAIC 262
 Qy 273 WTPRHVRDLFFSFVEETESLAAVFNLIHVSQVFFL---SSAVNPITYNLSRRRRAA 329
 Db 263 WLPRLHLFLIGSFQED-----YCHKRFIQOYVLAFLWLAASSTWYNPIIYCCLLRRRSG 317
 Qy 330 FRNVVSTPKWCHP-----RHRPQGPAPAKIIFLTECHLVLTEDAGPQFGSS 379
 Db 318 FR-LAFRCCHVMTPTEDKLELHTPSPS-----LRVNRCHTKELLMAGDVTSEAT 369

RESULT 8
 156444
 thyrotropin-releasing hormone receptor - mouse
 C/Species: Mus sp. (mouse)
 C/Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 11-Jan-2000
 C/Accession: I56444
 R/Seller, R.E.; Taylor, P.L.; Lamb, R.F.; Zabavnik, J.; Anderson, L.; Eldne, K.A.

J. Mol. Endocrinol. 10, 199-206, 1993
 A/Title: Functional expression and molecular characterization of the thyrotropin-rel
 A/Reference number: I56444; MUID:93249585
 A/Accession: I56444
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-411 <RES>
 A/Cross-references: GB:S60053; NID:g300151; PIDN:AAB26491.1; PID:g300152
 C/Superfamily: adenosine receptor A1

Query Match 19.4%; Score 403.5; DB 2; Length 411;
 Best Local Similarity 31.3%; Pred. No. 1.7e-26;
 Matches 117; Conservative 61; Mismatches 111; Indels 85; Gaps 13;

Qy 18 LNSTEEYLAHLCPKRDLSLPSVAVAL-----FLV-----GYMGLLCAVYVRH 65
 Db 9 LNDTE-----LPQVAVALEYQVVTLLVAVIGLIGVGNIMVVLVVRT 53
 Qy 66 QTLKPTNYTFLSLAVSDLLVL-LGMPL--EYEMMHVYPLFGVPGYFRTALPETV 121
 Db 54 KHMRTATNCTLVSLAVDMLVLAAGLPNTDSTYSGW-----YGVYGLCTITDYLG 108
 Qy 122 CFASILSVTVSERYVAIVHPRAKLESTRRRALRILSLWSEFVSFLPNTSINGIKF 181
 Db 109 INSSCSITFAFTIRYIAICHPKAQFLCTFSRAKKIIFVMAFTSIYCMLEFLDLNI 168
 Qy 182 QHPNGSSVGSATCTYTKPMVYNLIQATSLFYILPMTLSVLYLMGLR----- 235
 Db 169 STYKDAIV-----SCGKISIRNYSPITVLMDFGVYVMPMLATVLFARLFLNP 224
 Qy 236 -----KRDELEANKAVNIHR-----PSRKSVTKMLFVLVFAICWTPFH- 277
 Db 225 SDPRENSKTKMNDSTHONKNMNLTTNRCNSYSSKQVTKMLAVVYIIFALLMMPYRT 264
 Qy 278 --VDRLFES--EVEWTESLAAVFNLIHVSQVFFYLSAVNPITYNLSRRRRAFRNV 333
 Db 285 LVVNSFLSSPQENK-----FLFCRICIYLNLSAIPVYNLMSQKFAFR-- 332
 Qy 334 VSPTRCKWCHPRHR 347
 Db 333 -----KLCNCKQKP 341

RESULT 9
 S23436
 thyrotropin receptor - rat
 N/Alternate names: thyrotropin-releasing hormone receptor
 C/Species: Rattus norvegicus (Norway rat)
 C/Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 11-Jan-2000
 C/Accession: S23436; I53279; A49168; P00326
 R/de la Pena, P.; Delgado, L.M.; del Cantino, D.; Barros, F.
 Biochem. J. 284, 891-899, 1992
 A/Title: Cloning and expression of the thyrotropin-releasing hormone receptor from GH
 A/Reference number: S23436; MUID:92322017
 A/Accession: S23436
 A/Molecule type: mRNA
 A/Residues: 1412 <PEN>
 A/Cross-references: EMBL:X64630; NID:g57394; PIDN:CAA45913.1; PID:g57395
 R/Kimura, N.; Arai, K.; Sahara, Y.; Suzuki, H.; Kimura, N.
 Endocrinology 134, 432-440, 1994
 A/Title: Estradiol transcriptionally and posttranscriptionally up-regulates thyrotrop
 A/Reference number: I53279; MUID:94102223
 A/Accession: I53279
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-412 <RES>
 A/Cross-references: GB:D17469; NID:g464199; PIDN:BA04289.1; PID:g464200
 R/Zhao, D.; Yang, J.; Jones, K.E.; Gerald, C.; Suzuki, Y.; Hogan, P.G.; Chln, W.W.; T
 Endocrinology 130, 3529-3536, 1992
 A/Title: Molecular cloning of a complementary deoxyribonucleic acid encoding the thy
 A/Reference number: A49168; MUID:92283212
 A/Accession: A49168

A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-12, 'D', '14-290, 'K', '292-412 <ZHA>
 A:Experimental source: CH cells
 A:Note: Sequence extracted from NCBI backbone (NCBIN:104788, NCBI:P:104795)
 R:Yanada, M.; Monden, T.; Satoh, T.; Iizuka, M.; Murakami, M.; Iriuchijima, T.; Mori, M.
 Biochem. Biophys. Res. Commun. 184, 367-372, 1992
 A:Title: Differential regulation of thyrotropin-releasing hormone receptor mRNA levels
 A:Reference number: PQ0326; MUID:92231953
 A:Accession: PQ0326
 A:Molecule type: mRNA
 A:Residues: 30-58, 'P', '60-222, 'T', '224-261 <YAN>
 A:Experimental source: strain M1ster
 A:Note: the authors translated the codon ACA for residue 88 as Ala
 C:Superfamily: adenosine receptor A1
 C:Keywords: G protein-coupled receptor; transmembrane protein
 F:1-22/Domain: transmembrane #status predicted <TM1>
 F:32-54/Domain: transmembrane #status predicted <TM2>
 F:71-92/Domain: transmembrane #status predicted <TM3>
 F:116-140/Domain: transmembrane #status predicted <TM4>
 F:165-186/Domain: transmembrane #status predicted <TM5>

Query Match 19.4%; Score 403.5; DB 2; Length 412;
 Best Local Similarity 31.3%; Pred. No. 1.7e-26;
 Matches 117; Conservative 61; Mismatches 111; Indels 85; Gaps 13;

OY 18 LNSTEEYLAHLGPKRSDLSLPVSVAL-----IFLV-----GVMGNLLVCVYVRH 65
 DB 9 LNQTE-----LPPQVAVALGXOVVITLLVVICGIGVINIMVVLVVMRT 53
 OY 66 QTLKPTNYVLESLAVSDLVLL-LGMP-L--EIEEMHNHYFLPGPGCYEKTALFEIV 121
 DB 54 KIMRATTCYLSLAVADLVAVAGLPRTIDSIYGSW----VGYVGCCTITLQYIG 108
 OY 122 CFASILSVTVSVERVAVIHPFRALKLESTRRRLRLISLWSFSVSESLPMTSHIGIKF 181
 DB 109 INASSCSITAFIERIYAIICHPKIQOFLCTFSRAKKIIFVWAFISYICMLFPLDLNI 168
 OY 182 QHFPNGSSVPGSATCTVTKPMVNYLLIQAITSFLYIILPMTLSVLYIMGLRL----- 235
 DB 169 STYKDAIYI---SCGYKISRNYSPILMDGVYVPMILATLYGFIARILFLNPIR 224
 OY 236 -----KRDESLKANKVAVNIHR-----PSRSKVTMLFLVLYVFAICWTPFH- 277
 DB 225 SDPKNSNTWKWDSTHOKNNMLNTNRCFNSTVSSRKQVYMLAVVILFLMLMPTPT 284
 OY 278 --VDRLFPS--FVEEMTESLAVFNLIHVSGVFYLSASVNPPIIYNLSRRRAAFRNV 333
 DB 285 LVVNSFLSPQENM-----FLLFRCICILNSAINPIVYINLMSQKFRFAFR-- 332
 OY 334 VSPCKKCHPRHRP 347
 DB 333 -----KLCNCKOKP 341

RESULT 10
 S65693
 Opioid receptor mu variant MOR1A - human
 C:Species: Homo sapiens (man)
 C:Date: 12-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 05-Nov-1999
 C:Accession: S65693; S51216
 R:Bare, L.A.; Mansson, E.; Yang, D.
 submitted to the EMBL Data Library, July 1994
 A:Description: Expression of two variants of the human mu opioid receptor mRNA in SK-N-SH
 A:Reference number: S65693
 A:Accession: S65693
 A:Molecule type: mRNA
 A:Residues: 1-392 <BAR>
 A:Cross-references: EMBL:U12569; NID:9607911; PIDN:AAB60354.1; PID:9607912
 R:Bare, L.A.; Mansson, E.; Yang, D.
 FEBS Lett. 354, 213-216, 1994
 A:Title: Expression of two variants of the human mu opioid receptor mRNA in SK-N-SH cell

A:Reference number: S51215; MUID:95046336
 A:Accession: S51216
 A:Molecule type: mRNA
 A:Residues: 387-392 <BAN>
 C:Superfamily: vertebrate rhodopsin

Query Match 19.4%; Score 402.5; DB 2; Length 392;
 Best Local Similarity 30.4%; Pred. No. 2e-26;
 Matches 107; Conservative 65; Mismatches 133; Indels 47; Gaps 11;

OY 8 SWIHDPMLKYLNSTEYLAHLGPKRSDL-----SLPVSVALIFLV 50
 DB 30 SWVN-----LSHDGDLSDPCGPNRTIDGSDLSLCPPTGSPMTATITMALYSIVCV 83
 OY 51 GVMGNLLVCVYVRHQTLPYNYVLESLAVSDLVLLGMPLE-IYEMHNHYPLPGPV 109
 DB 84 GLEFNLVWYVIVRATKMTATNIYIFNLADALATST-LPFGSVNYLMGTWP--FGTI 140
 OY 110 GCFYFTALFEYVCFPSILSVTVSVERVAVIHPFRALKLESTRRRLRLISLWSFSVVF 169
 DB 141 LCKIVISIDYVMEFYSIFLCTMSVDRIYAVCHPAKALDFRPNRAKKIINVCNILLSSAI 200
 OY 170 SLPTSHIGIKFQHPNGSSVPGSATCTVT--KPMVNYLLIQAITSFLF-YILPMTLSV 226
 DB 201 GLPVMFINTATKQ-----GSIDCTITFSPHTWYEMNLKICVFIFAFIMPLVITV 252
 OY 227 LYLLGRLIKRDESLKANKVA-VNIHRPSKSVTKMLFLVLYVFAICWTPFHVDRLPFSF 285
 DB 253 CYGLMILRLKSVRLMSGSKEDRNLR---ITRMVLVAVVAVFCWTPPIHIVYIKAL 307
 OY 286 VEWETESLAAVNLHIVSGVFYLSAVNPPIIYNLSRRRAAFRNVSP 337
 DB 308 VTIPETFOYTSWHFCIALG---YTNSCNLPVLAFLDENFRCEFRCEPIPT 356

RESULT 11
 I56553
 N:Alternate names: MOR1 protein; opioid receptor mu
 C:Species: Homo sapiens (man)
 C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 19-May-2000
 C:Accession: I56553; A38991; S41075; S51215
 R:Meister, A.; Hurley, J.H.; Bye, L.S.; Campbell, A.D.; Chen, Y.; Tian, M.; Liu, J.; S
 J. Neurosci. 15, 2396-2406, 1995
 A:Title: The human mu opioid receptor: modulation of functional desensitization by ca
 A:Reference number: I56553
 A:Accession: I56553
 A:Status: nucleic acid sequence not shown; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-400 <RES>
 A:Cross-references: GB:L29301; NID:9459831; PIDN:AAA73958.1; PID:9459832
 R:Wang, J.B.; Johnson, P.S.; Persico, A.M.; Hawkins, A.L.; Griffin, C.A.; Uhl, G.R.
 submitted to GenBank, August 1994
 A:Reference number: A38991
 A:Accession: A38991
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-50, 'N', '52-233, 'V', '235-400 <MAN>
 A:Cross-references: GB:L25119; NID:9452072; PIDN:AAA20580.1; PID:9452073
 R:Wang, J.B.; Johnson, P.S.; Persico, A.M.; Hawkins, A.L.; Griffin, C.A.; Uhl, G.R.
 FEBS Lett. 338, 217-222, 1994
 A:Title: Human mu opiate receptor. cDNA and genomic clones. pharmacologic characteriz
 A:Reference number: S41075; MUID:94139928
 A:Accession: S41075
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-50, 'N', '52-400 <MA>
 R:Bare, L.A.; Mansson, E.; Yang, D.
 FEBS Lett. 354, 213-216, 1994
 A:Title: Expression of two variants of the human mu opioid receptor mRNA in SK-N-SH c
 A:Reference number: S51215; MUID:95046336
 A:Accession: S51215

us-09-609-146-25.rpr

A:Residues: 1-22, 'I', '24-293, 'F', '294-374, 'H', '376-398 <RRI>
C:Comment: The endogenous ligand of this receptor is neurokinin 2 (substance K or neurokinin B)
C:Genetics:
A:Gene: GDB:TAC2R
A:Cross-references: GDB:126367; OMIM:162321
A:Map position: 10q11.10q21
A:Introns: 131/2; 247/3; 313/2
C:Superfamily: neurokinin 1 receptor
C:Keywords: G protein-coupled receptor; glycoprotein; hormone receptor; transmembrane protein
F:33-59/Domain: transmembrane #status predicted <TM1>
F:70-93/Domain: transmembrane #status predicted <TM2>
F:111-119/Domain: transmembrane #status predicted <TM3>
F:150-169/Domain: transmembrane #status predicted <TM4>
F:194-222/Domain: transmembrane #status predicted <TM5>
F:253-275/Domain: transmembrane #status predicted <TM6>
F:288-310/Domain: transmembrane #status predicted <TM7>
F:11,19/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:106-181/Disulfide bonds: #status predicted

Query Match 19.0%, Score 394, DB 1, Length 398;
Best Local Similarity 28.9%, Pred. No. 1,16-25;
Matches 102; Conservative 64; Mismatches 153; Indels 34; Gaps 10;

OY 41 SVAVALIFVGMGLVLCMVIVRHQTLKPTPNYVLFSLASDLLVLLGLPLEYEMKH 100
DB 37 ANVALVALVAATGAAVIAVIMVILAHRRMRTVNYIVNILLADLCMAAFNFAFNVASH 96
OY 101 NVPFLGPGVGCYFKALFEVCFASILSVTVSVSEYVAIVHPRKALESTRRLRLILS 160
DB 97 NI-WFGNAFCFYQLPFTAMFVSIYSMTAILADRYAIVHPQPRLSAPSTKA--VIA 153
OY 161 LVMSESVVFSLEPNTSIHGKIFQHPFGSSVPGSATCTVT-----KPMVYNLLIQATS 213
DB 154 GIMVALVALASPOCFYSTVTMDQ-----GATCCVAMPEDSGKTLLEHLVIA-- 203
OY 214 FLFLYLLPMTLISVLYLGLGLRLKRDSELEANKVAVNI--HRSRKSVTKMLFVLVYFAIC 272
DB 204 -LIVFLPLVAVMEVAASVIGITLMRAVPGHOAHGANGLHRLKKEFVTWMLVLYTFEAC 262
OY 273 WTPFHVDRLFESFVEWKESSLAAVFNLIHVVSGVFYFL---SSAVNPITVNLSSRFAA 329
DB 263 WLPYHLYPLILSGFOED-----LYCHKPFIQOYVALFVWLMSSTMNPITTYCCLNIRFSG 317
OY 330 FRNVVSPICKWCCHPRHPO--GPPAQKIIFLTECHLVLETEADGAPPGQSS 379
DB 318 FR-LAFRCRPMVTPYKEDKLELTPTSLSTRVNRCHTKETLEFMAGDTAPSEAT 369

RESULT 14
S55524
neurokinin 3 receptor - mouse (fragment)
N:Alternate names: neuromedin K receptor; NK-3 receptor
C:Species: Mus musculus (house mouse)
C:Date: 01-Aug-1995 #sequence_revision 01-Sep-1995 #text_change 19-May-2000
C:Accession: S55524; I73045
R:Maroteaux, L.
submitted to the EMBL Data Library, June 1995
A:Reference number: S55524
A:Accession: S55524
A:Molecule type: mRNA
A:Residues: 1-385 <MA>
A:Cross-references: EMBL:X87823; NID:9861055; PIDN:CAAG1088.1; PID:9861056
R:COOK, G.A.; ELLIOTT, D.; METWELL, A.; BLUM, A.M.; SANDOR, M.; LYNCH, R.; WEINSTOCK, J.
J. Immunol. 152, 1830-1835, 1994
A:Title: Molecular evidence that granuloma T lymphocytes in murine schistosomiasis manse
A:Reference number: 156216; MUID:94165478
A:Accession: I73045
A:Status: Preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 103-197, 'S', '199-266, 'P', '268-328 <COO>
A:Cross-references: GB:I27827; NID:9450288; PIDN:AAAI7893.1; PID:9480780
C:Superfamily: neurokinin 1 receptor

[illegible]

C:Keywords: alternative splicing; G protein-coupled receptor; glycoprotein; transmembran

Query Match 18.6%; Score 387; DB 2; Length 398;

Best Local Similarity 29.2%; Pred. No. 4.1e-25; Matches 108; Conservative 66; Mismatches 134; Indels 62; Gaps 13;

```

OY 6 NASMIDPLM-----KYLNSTEEYLAHL-----CGPKRSDL----- 36
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 9 NISDCSDPLAPASCSPAPGSMLN-----LSHVGNQSDPCGPNRTGLGSHSLCPQTGSP 63
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 37 ----SLPVSVAVALIFLVGVGNLVCMTVIRHQTLKTPPTNYLFLSLAVSDDLVLGMP 92
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 64 SMTATITMALYSIVCVGLFNGFLVMYIVRYTKMKATATNIYIFNLADALATSP-LP 122
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 93 LE-IYEMMHNYPLFGVGCYEKTALEETVCEASILSVTVSERYVAIVHPRAKLEST 151
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 123 FOSVNTLMGTWP--FGNLLCKIVISIDYNNMFTSIFTLCTMSVDRYIACHPVKALDERT 180
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 152 RRRALRLSLVMSFSVFSLPNTSIHQIKPQHFPNGSSVPSATCTVT--KPMVYNYLI 209
    | | : | : | | | | | : : : | | | | | | | | | | | | | |
DB 181 PRNAKIVNVCNMILSSAIGLPVPMFMAITTKYRQ-----GSIDCTLTFSHTWYENL 232
    | | : | : | | | | | : : : | | | | | | | | | | | | | |
OY 210 QATSFLE-YLPMTLISVLYYLMGLRLKRDSELEANKVA-VNIHRSRKSVTKMLEVYL 267
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 233 KICVFTEAFIMPVLIITVCYGLMILRLKSVMLSGSKEDRNLR-----ITRMVLVYVA 287
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 268 VFAICWTPFHVDLFFSFVEWETESLAAVENLIHVVGCVFFYLSAVNPITLYNLSRRR 327
    | | : | | | | : : : | | : : : | | : : : | | : : : | | :
DB 288 VFIVCWTPHIHVIYIKALITIPETTFOTVSMHFCIALG---YTNSCUNPVLYAFLDENFK 344
    | | : | | | | : : : | | : : : | | : : : | | : : : | | :
OY 328 AAFRNVSPT 337
    | | | |
DB 345 RCFREFCIPT 354
    | | | |

```

Search completed: April 22, 2002, 14:17:44
Job time: 297 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OH nucleic - nucleic search, using sw model

Run on: Apr11 22, 2002, 17:16:21 ; Search time 2258.39 seconds
(without alignments)
5857.294 Million cell updates/sec

Title: US-09-609-146-24
Perfect score: 1231
Sequence: 1 gttgtgatttaagtcag.....aggagtggtcagaagcctc 1231

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues
Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: em_estfun:*
2: em_esthum:*
3: em_estlin:*
4: em_estom:*
5: em_estcpl:*
6: em_estlba:*
7: em_estlro:*
8: em_estlov:*
9: em_hlc:*
10: gb_est1:*
11: gb_est2:*
12: gb_hlc:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_fod:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	281.8	22.9	555	10	BE751626 203809 MA
2	214.2	17.4	465	10	BE751310 203132 MA
3	187.8	15.3	941	13	CNS036WH
4	179.2	14.6	929	11	BE871983 60148032
5	139	11.3	414	11	M45474
6	114	9.3	223	10	BS587450
7	96.6	7.8	1079	13	CNS04P4T
8	86.4	7.0	389	10	AA562357
9	86.4	7.0	450	11	R13850
10	86.4	7.0	500	11	BF603623
11	85.8	7.0	959	13	CNS04GXP
12	83.6	6.8	448	11	BF323227

c	13	82	6.7	480	13	A2616812
c	14	77.8	6.3	379	11	H11359
c	15	77.8	6.3	817	13	CNS01X12
c	16	74.6	6.1	835	13	CNS020N1
c	17	72.2	5.9	471	10	AU204159
c	18	70.8	5.8	741	11	B1224313
c	19	69	5.6	1189	12	AK005368
c	20	68.8	5.6	634	11	BG714306
c	21	67.4	5.5	975	13	CNS039RV
c	22	67.4	5.5	1041	11	BG293064
c	23	66	5.4	612	11	BF446062
c	24	66	5.4	825	10	AU079556
c	25	63.6	5.2	396	10	BE654108
c	26	62.8	5.1	979	13	CNS048AL
c	27	62.6	5.1	875	10	AL537553
c	28	62.4	5.1	563	13	A2357931
c	29	62	5.0	460	10	AW206493
c	30	61.2	5.0	662	10	A1820631
c	31	61	5.0	788	10	AL541044
c	32	60.4	4.9	462	10	AM138148
c	33	60.4	4.9	544	10	AT344017
c	34	60.4	4.9	549	11	BE938149
c	35	60.2	4.9	638	10	AM139137
c	36	60	4.9	566	10	BE266564
c	37	60	4.9	988	13	CNS018EH
c	38	59.6	4.8	335	13	BH034857
c	39	59.4	4.8	753	10	AT692402
c	40	59.2	4.8	768	10	BE746231
c	41	59	4.8	523	11	BE437994
c	42	59	4.8	861	11	BE798784
c	43	58	4.7	339	13	AO015065
c	44	58	4.7	519	13	AO019411
c	45	58	4.7	597	10	AT693287

ALIGNMENTS

RESULT 1
LOCUS BE751626 555 bp mRNA
DEFINITION 203809 MARC 280V Bos taurus cDNA 5', mRNA sequence.
ACCESSION BE751626
VERSION BE751626.1 GI:10165618
KEYWORDS EST.
SOURCE COW.
ORGANISM Bos taurus

REFERENCE
AUTHORS Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T., Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G., Pereira,G., Holt,I., Karamecheva,S., Liang,F., Quackenbush,J. and Keeler,J.W.
1 (bases 1 to 555)

Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle
Genome Res. 11 (4), 626-630 (2001)
21180013
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@mail.marc.usda.gov
Single pass sequencing. Bases called and all trimmed with phred v0.980904.e. Vector identified by cross_match with the -m1nscore 18 and -mismatch 12 options.

PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCACATCAGCAGC
Plate: 44 row: O column: 6

[illegible]

Db	150	CTCTGCTGTGGCTGCTGCTGCATGCTGCTGCGGGGCGGCCAACACACAGCGTGCAGGCATC	91
QY	560	aagttccagcac	571
		:	
Db	90	GAGGGSYGGCGC	79
RESULT	4		
LOCUS	BE871983		
DEFINITION	BE871983	929 bp	mrna
ACCESSION	601448032f1	NIH_MGC_65	Homo sapiens cdna clone IMAGE:3852151 5',
VERSION	BE871983		
KEYWORDS	EST.		
SOURCE	BE871983.1	GI:10320759	
ORGANISM	human.		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
TITLE	NIH-MGC http://mhc.nci.nih.gov/		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgapsb@mail.nih.gov		
	Tissue Procurement: ATCC		
	CDNA Library Preparation: Life Technologies, Inc.		
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LML)		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LML at:		
	http://image.llnl.gov		
	Plate: FLAM9573 row: P column: 08		
	High quality sequence start: 11		
	High quality sequence stop: 602.		
FEATURES	Location/Qualifiers		
source	1..929		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone_1ib="IMAGE:3852151"		
	/clone_1ib="NIH_MGC_65"		
	/tissue_type="adenocarcinoma"		
	/lab_host="DH10B (phage-resistant)"		
	/note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;		
	Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT		
	Average insert size 1.8 kb. Library constructed by Life		
	Technologies.		
BASE COUNT	186 a	252 c	263 t
ORIGIN			
Query Match	14.6%	Score 179.2;	DB 11; Length 929;
Best Local Similarity	67.7%	Pred. No. 7.5e-37;	
Matches 394; Conservative	0;	Mismatches 158;	Indels 30; Gaps 9
QY	3	tgtgatcttaagctcagtaatggaagaactgaaatgctcctgataccac-----	55
Db	123	TGTGATTTTAAATGTCTCAGGATGGAAAAATCTCAAAATGCTTCCGTGATCTACACGAGAA	182
QY	56	-----gactcctcctgaagtaacttgacagacagacagagaagtaacttggccca-cctgt	107
Db	183	ACTACAAAGATCCATTCCAAAAMACCTGGAACAGACCGAGGAGTATCTGGCCTTCCCTCT	242
QY	108	gtgacgaagaagcagagagactatccctcccgltgctgtgacatgacgtcgtgacttcc	167
Db	243	GCGGACCTCGGCGAGGACCACTTCTCTCCCGCTGCTGTGTGTATGTGCCAATTATGG	302
QY	168	tgtgtgaggtgaatgagcgaatctctgtgtgtgcacgtgtgacatccagacttga	227
Db	303	TGGGGGGGGGTCATGCGCAATGCTCTGCTGCTGCTGTGATCTGACGACACAGCGCTATGA	362
QY	228	agacaccaccacactatcattcattcagcttggcag---tctcagatcgtcgtgctcgc	284

Db	Accession	Version	KeyWords	Source	Organism	Title	Journal	Comment
Db	363	AGACC	CCACGAA--TACTACCTCTTCAGCCTTGCCGGGTCTGACCTCTCGGGTCTTGC					
Qy	285	tcttgaggatgctcttggaatct--acagatgctggacaataacccttctcttcg						
Db	422	TCCTTTGGAAAGGCCCGGTGAGGTCTAATGACATGTGGGGCAACTACCCCTTTTGGGTC						
Qy	342	gg---ctgtggatgctactccaagacagccctcttgagacgtgtgtcttgcctca						
Db	482	GGGCCCGTTGGGGCTGTACATTCAGAGAGGCCCTTTTTCAGACCGGTTCCTCGCTTCA						
Qy	399	tctctagtgctcaacggttagcgtatagcgtctatgtgtgcaattgcaacccttcacg						
Db	542	TCTCTAGATCACACCG--TCAGGTGAGCGCCTACGTGGCAATCTTACACCCGTT--CGG						
Qy	459	ccaagctggagagacagcgcgacgagcgccctcagatccctcagcctagtctgagcttct						
Db	599	CCAAATTGAGGAGACACCGCGGTGCGGGCCTCAGGAATCTCGGAATTCTCTGGGGTACT						
Qy	519	ctgtgtcttcttctcttgcccaatcaccagcatccatgacatca						
Db	659	CCGACACTCTTCCTCG--CAAAACGACGATCATGTGGCTTCA						
RESULT	5							
LOCUS	N45474	414 bp	mRNA	EST	13-FEB-1996			
DEFINITION	yys9b04.r1 Soares_multiple_sclerosis_2NDHMS							
ACCESSION	N45474							
VERSION	N45474.1	GI:1186640						
KEYWORDS	EST.							
SOURCE	human.							
ORGANISM	Homo sapiens							
REFERENCE	1 (bases 1 to 414)							
AUTHORS	Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Holtzman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevisan, E., Waterston, R., Williamson, A., Wohlmann, P., and Wilson, R.							
TITLE	The WashU-Merck EST Project							
JOURNAL	Unpublished (1995)							
COMMENT	Contact: Wilson RK Washington University School of Medicine 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@wustl.wustl.edu This clone is available royalty-free through LNL; contact the Image Consortium (info@image.llnl.gov) for further information. Seq primer: 17 High quality sequence stop: 264.							
FEATURES	source							
	1..414							
	/organism="Homo sapiens"							
	/db_xref="GDB:3896183"							
	/db_xref="taxon:9606"							
	/clone="IMAGE:277807"							
	/clone_1ib="Soares_multiple_sclerosis_2NDHMS"							
	/sex="male"							
	/tissue_type="multiple sclerosis lesions"							
	/dev_stage="Age 46"							
	/lab_host="DH10B (ampicillin resistant)"							
	/note="Vector: pRT3D (Pharmacia) with a modified polylinker V-TPRE; phagemid; site_1: Not I - site_2: Eco R ; 1st strand cDNA was primed with a Not I - oligo(dT) primer 15. TGTATCAATCTGAAGTGGAGCGGCCGCAATTTTTTTTTTTTTTTT 3'). double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pRT3 vector (Pharmacia). Library went through one round of normalization to a Col = 5. Library constructed by Bento							

	Soares and M.Fatima Bonaldo.	RNA from 4 multiple sclerosis
	lesions from one patient was kindly provided by Dr. Kevin G. Becker (NINDS/NIH).	"
	Query Match	11.3%, Score 139; DB 11; Length 414;
	Best Local Similarity	73.3%; Pred. No. 2.7e-26;
	Matches 178; Conservative	0; Mismatches 65; Indels 0; Gaps 0;
BASE COUNT	113 a 123 c 71 g 105 t	2 others
ORIGIN		
OY	ccatatctataactcctctggttcgcggcgcttccaggcgagcctttgaatagtgtctccc	1028
DB	1 CCATTATCTATAACTGTTCTGTCCGCCGTCACAGCAGCATTCCAGAATGGATCTCTT	60
OY	ctacctgcacaattggtgcccatccccgcgatccgacagaagaccctcagaccgaagaatc	1088
DB	61 CTTTCCACACAACAGTGCGCACTCTCCAGCATGACCACAGTTGCACACTGCCCCAGCGGAACA	120
OY	tctcttgacagaaatgtaactcctgtagagcgtgcacagagatgacagagcccccaagtccctg	1148
DB	121 TCTTCTGTGACAAATGCCACTTTTGTHGAGAGCTGACCGAAGATATAGTGTCGCCAATTCCC	180
OY	gtcacatcatccatccacaacacccaacttacacagggccccctgtgcagagagaggtaccat	1208
DB	181 GTCACTGATCATCATGACACAACATCTCACCTCCCAAACGCCCTCTGTAGTAGAACAGATGTCAA	240
OY	1209 aaa 1211	
DB	241 GAA 243	
RESULT		
BBS87450	223 bp	mRNA EST 30-NOV-2000
LOCUS	BBS87450 RIKEN full-length enriched, adult female vagina Mus	
DEFINITION	musculus CDNA clone J93000IG13 5', mRNA sequence.	
ACCESSION	BBS87450	
VERSION	BBS87450.1 GI:11483994	
KEYWORDS	EST,	
SOURCE	house mouse,	
ORGANISM	Mus musculus	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 223)	
AUTHORS	Alizawa,K., Akahira,S., Akimura,T., Arai,A., Arakawa,T., Carninci,P., Hasegawa,K.T., Hayatsu,N., Hirooka,T., Hiozane,T., Hodoyama,Y., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kawai,J., Kojima,Y., Konno, H., Kusabade,M., Matsuyama,T., Miyazaki,A., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T., Owa,C., Sakai,C., Sakai,K., Sasaki,D., Sato,K., Shibata,K., Shibata,Y., Shinogawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Tanaka T., Toya,T., Watabiki,A., Yamamuro,T., Yasunishi,A., Yoshida,K., Yoshikita,A., Muramatsu,M. and Hayashizaki,Y.	
TITLE	RIKEN Mouse ESTs (Alizawa,K. et al. 2000)	
JOURNAL COMMENT	Unpublished (2000) Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suenitcho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel.: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@gsr.riken.go.jp/ URL:http://genome.gsc.riken.go.jp/	
Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoke,S., Sasaki, N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y. Thermotabilization and thermostable enzymes by thermolase and its application for the synthesis of full length cdna. Proc. Natl. Acad. Sci. U.S.A. 95 (2). 520-524 (1998) Itoh.M., Katsunami,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J., Tomaru,Y., Carninci,P., Shibata,T., Ozawa,T., Muramatsu,M., Okazaki, Y. and Hayashizaki,Y.		

AUTHORS
Roest-Crolius, H., Jalllon, O., Dasilva, C., Fitzames, C., Fisher, C.,

TITLE	Waterston, R. The Washu-HHMI Mouse EST Project
-------	---

TITLE	Keele, J.W.
JOURNAL	Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle
MEDLINE	Genome Res. 11 (4), 626-630 (2001)
COMMENT	21180013
	Contact: Smith TPL

Fax: 402 762 4390
Email: smith@emall.maricopa.usda.gov
Single pass sequencing. Bases called and alt-trimmed with phred
v0.980904.e. Vector identified by cross-match with the -mismatches
-mismatch 12 options.

FEATURES

FORWARD:	AGGAACACGCTATGACCAT
BACKWARD:	GTTTCCCACTACAGAC
Plate:	49
row:	I
column:	6

Seq primer: ATTTAGGTGACACTTATG

Location/Qualifiers

```

source
1..500
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_11b="MARC 3BOV"
/tissue_type="pooled"
/lab_host="DH10B"
/notice="Vector: PCMV SPOR6; Site.1: XbaI; Site.2: XhoI;
Library made from pooled tissue from marrow, alveolar
macrophage, ovary, fetal semitendinosus muscle, and fetal
longissimus muscle."
BASE COUNT
71 a 179 c 139 g 111 t

```

Query Match	7.08;	Score	86.4;	DB	11,	Length	500;
Best Local Similarity	52.88;	Pred.	NO.	2.7e-12;			
Matches	211;	Conservative	0;	Mismatches	186;	Indels	3;
						Gaps	1

[illegible]

RESULT	11
CNS04GXP	
LOCUS	959 bp DNA
DEFINITION	Tetradon nigroviridis genome survey sequence PIC-ori end of clone 10K06 of library G from Tetradon nigroviridis, genomic survey sequence.

ACCESSION	AL290230
VERSION	AL290230.1
KEYWORDS	GI:8028810
SOURCE	GSS: genome survey sequence.
ORGANISM	Tetraodon nigroviridis.
	Tetraodon nigroviridis
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Acanthopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodon.
REFERENCE	1 (bases 1 to 959)
AUTHORS	Roeest,Crollius,H., Jalllon,O., Dasilva,C., Fitzames,C., Fisher,C., Bouneau,T., Billault,A., Quetler,F., Saurin,W., Bernot,A. and Weissenbach,J.
TITLE	Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 959)
AUTHORS	Roeest,Crollius,H., Jalllon,O., Dasilva,C., Bouneau,T., Fisher,C., Bernot,A., Fitzames,C., Mincker,P., Brotlier,P., Quetler,F., Saurin,W. and Weissenbach,J.
TITLE	Human gene number estimate provided by genome wide analysis using Tetraodon nigroviridis DNA sequence
JOURNAL	Unpublished
REFERENCE	3 (bases 1 to 959)
AUTHORS	Genoscope.
TITLE	Direct Submission
JOURNAL	Submitted (12-APR-2000) to the EMBL/GenBank/DBD databases
COMMENT	This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/tetraodon .

FEATURES	source
location/Qualifiers	1. 959
/organism="retrovirus nairovirus"	
/db_xref="taxon:99883"	
/clone="109x06"	
/clone_1b="G"	
/note="Genosense sequence ID : C0BC109BF03SP1-end ; puc-011"	
BASE COUNT	184 a 293 c 290 g 184 t 8 others
ORIGIN	

Query Match	7.0%	Score 85.8	DB 13	Length 959
Best Local Similarity	55.4%	Pred. NO. 4.8e-12		
Matches 181; Conservative		3; Mismatches 140;	Indels 3;	Gaps 1.

[illegible]

QY	454	CGGAGCCAAGCTGTGAGAGCAAGCGGCG	480
Db	708	CCGCTACAAAGGCCCTGGGCGGGAACG	734
RESULT	12		

[illegible]

Query Match	5.7%	Score 82:	DB 13:	Length 480:
Best Local Similarity	58.8%	Pred. No. 3.9e-11:		
Matches 161:	Conservative	0:	Mismatches 110:	Indels 3:
			Gaps	1:


```

Qy 810 tctctctgctcgtgtgttgcacatcgtcgtaccctccatctgaccggtctctc 869
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 274 TTGACCTGTTGTGTATTCGGCATCTGCGGCTCCATTCATCTGACCGTATCATGT 215

Qy 870 tcaagctgtggaag--agtggacagagtcctcgtcgtgtgtgtcaacctcaccatg 926
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 214 GGAGCTGCTGTATGACACATCAACGAAGGCTTGACCTGGCTGACAGTGTGTCCACA 155

Qy 927 tgtatcagagtgctctcttctatcagctcgggtcaacccctatctataaccctcc 986
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 154 TTGCTCTGGCAATCTTCTTATCTCGGCTCAGCAGCAACCCGGTGGCTTACACCTCA 95

Qy 987 tctcagcgctcctcggcgcccttcgaatgtgtctccctaccctcgaaatgtgcc 1046
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 94 TGTCTACCTCGCTCCGACAGACACTTCCGCAAGCCCTGGCGCTTGGAACCCAGTCTGTGC 35

Qy 1047 atccccggacatcgccacagggacctccagccca 1080
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 34 ATCGCCGCCAACCCATCATGCTTCCCATGCGA 1

RESULT 14
LOCUS H11359 379 bp mRNA EST 26-JUN-1995
DEFINITION ym13d04.f1 Soares infant brain INIB Homo sapiens cDNA clone
IMAGE:47842 5' similar to contains 11 repetitive element ; , mRNA
sequence.
ACCESSION H11359
VERSION H11359.1 GI:876179
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 379)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
M., Hulman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
Ridkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
R., Williamson,A., Wohlmann,P. and Wilson,R.
The Washu-Merck EST Project
Unpublished (1995)
JOURNAL Washington University School of Medicine
COMMENT Contact: Wilson RK
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 3363
High quality sequence stops: 321
Source: IMAGE Consortium, LNLN.
This clone is available royalty-free through LNLN ; contact the
IMAGE Consortium (Info@image.lnl.gov) for further information.
Insert length: 3363 Std Error: 0.00
Seq primer: M13Rev
High quality sequence stop: 321.
Location/Qualifiers
1. 379
/organism="Homo sapiens"
/db_xref="GDB:420383"
/db_xref="taxon:9606"
/clone="IMAGE:47842"
/clone.lib="Soares infant brain INIB"
/sex="female"
/dev_stage="73 days post natal"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: whole brain; Vector: Lambda BA; Site_1: Not
I; Site_2: Hind III; 1st strand cDNA was primed with a Not
I - oligo(dT) primer [5';
AACGGAAGAAATTCGCGCCGACAGAAATTTTCTTTTCTTTT 3'];
double-stranded cDNA was ligated to Hind III adaptors
(Pharmacia), digested with Not I and directionally cloned
into the Not I and Hind III sites of the Lambda BA vector.

```

```

BASE COUNT 124 a 63 c 74 g 115 t 3 others
ORIGIN

Query Match 6.3%; Score 77.8; DB 11; Length 379;
Best Local Similarity 79.8%; Pred. No. 4.7e-10;
Matches 91; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Qy 722 aggcgtgaagagatgaatcccttgaggcgaacaaagtgtgtgaatactcagacc 781
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 266 AGNCTAAAGAAACAAATCTCTGAGCGAGATGAGAGGAATATTCATAAGACCC 325

Qy 782 tctagaagtcgtaccagaatgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 835
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 326 TCCAGAAATATCATGACACAGATGCTGTGTGTGTGTGTGTGTGTGTGTGTATC 379

RESULT 15
LOCUS CNS01XJ2 817 bp DNA GSS 12-MAY-2000
DEFINITION Tetradon nigriviridis genome survey sequence PUC-Or1 end of clone
202F23 of library G from Tetradon nigriviridis, genomic survey
sequence.
ACCESSION AL171767.1 GI:7809824
VERSION AL171767
KEYWORDS GSS; genome survey sequence.
SOURCE Tetradon nigriviridis.
ORGANISM Tetradon nigriviridis.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodon.
REFERENCE 1 (bases 1 to 817)
AUTHORS Roest-Crolius,H., Jallion,O., Dasilva,C., Fizames,C., Fisher,C.,
Bonneau,L., Billault,A., Quetler,F., Saurin,W., Bernot,A. and
Weissenbach,J.
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigriviridis
Unpublished
2 (bases 1 to 817)
REFERENCE Roest-Crolius,H., Jallion,O., Dasilva,C., Bonneau,L., Fisher,C.,
Bernot,A., Fizames,C., Wincker,P., Brotlier,P., Quetler,F.,
Saurin,W. and Weissenbach,J.
Human gene number estimate provided by genome wide analysis using
Tetraodon nigriviridis DNA sequence
Unpublished
3 (bases 1 to 817)
JOURNAL Direct Submission
REFERENCE Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases
AUTHORS This sequence is a single read and was generated as part of a large
JOURNAL scale clone-and-sequencing project of the Tetraodon nigriviridis
COMMENT genome. For more information, please take a look at
http://www.genoscope.cns.fr/Tetraodon.
Location/Qualifiers
1. 817
/organism="Tetraodon nigriviridis"
/db_xref="taxon:99883"
/clone="202F23"
/clone.lib="G"
/note="Genoscope sequence ID : C0AG202C12SP1-end :
PUC-Or1"
BASE COUNT 158 a 237 c 258 g 159 t 5 others
ORIGIN

Query Match 6.3%; Score 77.8; DB 13; Length 817;
Best Local Similarity 56.3%; Pred. No. 6e-10;
Matches 184; Conservative 1; Mismatches 138; Indels 4; Gaps 2;

Qy 157 gctgctctccctggtggggtatggtgacatctctcgtgtgcatgtgtgtccgaca 216

```

```
Db 492 GCTGACGGGAATCGTGGGGAACGGGTGACATCAGGTGAGCCAGGTGCTGAAAGCAGAA 433
Oy 217 tcaagacttgaagacaccccaactactatctctcaagcttggcaagtctcagatctgct 276
Db 432 CCGCTACCTGCAGAAAGCAAGTCACCGACCAATGATCAGCTGGCCCTGCTCCGACCTCCT 373
Oy 277 ggtcctgctcttgggagatgctcttggaaactta---cgaagatgagacaaattaccctt 333
Db 372 GGTGCTGCTCATCGGATGCGGATGAGCTTACAGCGCCATCTGTCCTTCACTTC 313
Oy 334 cctgtcgggacctgtggaatgctactcaagaacagccctcttggagactgtgtcttgc 393
Db 312 GCGTCCGGAACATCTCTTAAGATCTACAACTTCTGTTCGAGGC-GTGCAGTAAGC 254
Oy 394 ctccattctcaatgtlcaacaggttagcgtagagcgctatgtggcaatgtlcaaccctt 453
Db 253 CACCATTTGTAACATAGCCACGCTTAGCTTGAGCGCTAGGCGCATCTGCCACCCCTT 194
Oy 454 ccgagccaagctggaagacagcgcgcg 480
Db 193 CCGCTACAAAGGCCCTGGCGGGAAGCG 167
```

Search completed: April 22, 2002, 17:16:30
Job time: 10964 sec

Pred. NO. Is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

Result	No.	Score	Query	Match	Length	DB	ID	Description
C	1	1200.4	97.5	134	10	AB041229	AB041229 Rattus no	
	2	1186.4	96.4	1188	10	AF242875	AF242875 Rattus no	
	3	799.6	65.0	1298	6	AF272363	AF272363 Homo sapi	
	4	799.6	65.0	1658	6	AX109234	AX109234 Sequence	
	5	799.6	65.0	1658	6	AX109236	AX109236 Sequence	
	6	796.4	64.7	1658	6	AX109238	AX109238 Sequence	
	7	796.4	64.7	1658	6	AX109240	AX109240 Sequence	
	8	790.8	64.2	1248	9	AB041228	AB041228 Homo sapi	
	9	790.8	64.2	1248	9	AF292402	AF292402 Homo sapi	
	10	788	64.0	1239	9	AF242874	AF242874 Homo sapi	
C	11	677.6	55.0	1594	6	AX109242	AX109242 Sequence	
	12	677.6	55.0	1594	6	AX109244	AX109244 Sequence	
	13	481.4	39.1	215088	2	AC008571	AC008571 Homo sapi	
	14	472.6	38.4	801	6	AX147772	AX147772 Sequence	
	15	469.8	38.2	739	6	AF272362	AF272362 Homo sapi	
	16	330.6	26.9	1318	9	AF272362	AF272362 Homo sapi	
	17	301.4	24.5	1209	10	AB248873	AB248873 Rattus no	
	18	299.8	24.4	1239	10	AB038649	AB038649 Rattus no	
	19	285.6	23.2	1218	10	AF044602	AF044602 Mus muscu	
	20	266	21.6	828	9	HS0GPER1	AF044602 Homo sapi	
C	21	266	21.6	168880	9	AC0017104	AF0017104 Homo sapi	
	22	151.2	12.3	75950	2	AC073449	AC073449 Homo sapi	
	23	141	11.5	1098	6	AX154586	AX154586 Sequence	
	24	133.2	10.8	1092	6	AR156357	AR156357 Sequence	
	25	133.2	10.8	1350	10	AB001982	AB001982 Rattus no	
	26	132.8	10.8	1088	6	AR156353	AR156353 Sequence	
	27	132.8	10.8	1101	6	AX154584	AX154584 Sequence	
	28	132.8	10.8	1101	6	HSU60179	U601179 Human growt	
	29	128	10.4	1063	6	AR156351	AR156351 Sequence	
	30	128	10.4	1101	6	ASU60178	U601178 Sus scrofa	
C	31	114.2	9.3	1676	5	AF082210	AF082210 Spheroeoid	
	32	111.4	9.0	121652	34	AC086191	AC086191 Drosophill	
	33	111.4	9.0	140838	2	AC018176	AC018176 Drosophill	
	34	111.4	9.0	195212	3	AC007441	AC007441 Drosophill	
	35	111.4	9.0	225374	3	AE003703	AE003703 Drosophill	
	36	109.6	8.9	384	4	HS0GPER2	AF044601 Homo sapi	
	37	104.8	8.5	1161	6	AX154591	AX154591 Sequence	
	38	104.8	8.5	1239	6	AX154589	AX154589 Sequence	
	39	104.8	8.5	2040	9	AF034632	AF034632 Homo sapi	
	40	104.8	8.5	163284	9	AL137000	AL137000 Human DNA	
C	41	104.8	8.5	341560	2	AL1596304	AL1596304 Homo sapi	
	42	103.4	8.4	836	6	AR156355	AR156355 Sequence	
	43	102.2	8.3	3129	6	AR156356	AR156356 Sequence	
	44	102.2	8.3	3129	10	RNU94321	U94321 Rattus norv	
	45	100.2	8.1	870	9	HS060181	U60181 Human growt	

RESULT	1
LOCUS	AB041229
DEFINITION	Rattus norvegicus mRNA for G protein-coupled receptor TGR-1, complete cds.
ACCESSION	AB041229.1 GI:10257382
VERSION	G protein-coupled receptor TGR-1.
KEYWORDS	Rattus norvegicus cDNA to mRNA.
SOURCE	Rattus norvegicus
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE	1 (sites)
AUTHORS	Hosoya,M., Moriya,T., Kawamata,Y., Ohkubo,S., Fujii,R., Matsui,H., Shikama,I.Y., Fukusumi,S., Habata,Y., Hinuma,S., Onda,H., Nishimura,O. and Fujino,M. Identification and Functional Characterization of a Novel Subtype of Neuregulin Receptor
TITLE	J. Biol. Chem. 275 (38), 29528-29532 (2000)
JOURNAL	

PUBMED 10887190
 2 (bases 1 to 1314)
 Fujii, R., Shintani, Y. and Hinuma, S.
 TITLE Direct Submission
 JOURNAL Submitted (03-APR-2000) to the DDBJ/EMBL/GenBank databases. Shuji
 Hinuma, Takeda Chemical Industries, LTD, Discovery Research
 Laboratories 1, Wadai 10, Tsukuba, Ibaraki 300-4293, Japan
 (E-mail: Hinuma_Shuji@takeda.co.jp, Tel: 81-298-64-5035,
 Fax: 81-298-64-5000)
 FEATURES
 source location/Qualifiers
 1. 1314
 /organism="Rattus norvegicus"
 /db_xref="taxon:10116"
 gene 127..1314
 /gene="TGR-1"
 127..1314
 /gene="TGR-1"
 /codon_start=1
 /product="G-protein-coupled receptor TGR-1"
 /protein_id="BAB13722.1"
 /db_xref="GI:10257383"
 /translation="MGKLENSMIDPLMKYINTEEYLAHCGPKRDLSPVAV
 ALIFVGWGNLNCMVIVRHOTLTPRYVLESLASDLILLCMPLEIEMHNY
 PLFEPVGCYFTALPEYICFASISVTIVSERYVAIYHPRAKLESTRRALRLIS
 LWSFSSVFSPLPNTSIHGKIFQHPGSSVPSACVYKPMVYNLLIQATSFLEYI
 LPMTLISVLYLMDGLRKDESLKANKVAVNTHRSKSVTKMLFVLYVFAICMTPF
 HYDRLEFSEVEMTESLAAVFNLIHVSVFYLSSAVNPDIYNLSRFRRAFRNV
 SPTCKMCHPRHQPGCPAPAKIIFLTRECHMELTEDAGPQPGQSSIHNNLTMAPCAG
 EYP"
 BASE COUNT 269 a 397 c 305 g 343 t
 ORIGIN
 Query Match 97.5%; Score 1200.4; DB 10; Length 1314;
 Best Local Similarity 99.5%; Pred. No. 0;
 Matches 1204; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 Oy 1 gtcttgatcttcaagctcaagtaattggaagaacttgaagaacttcttgatccagatcc 60
 Db 105 GTTGTGATTTTAAAGCTCAATATGGGAAACCTTGAAGAACTTCCTGATCCACGATCC 164
 Oy 61 tctcatgaagtaacttgaaacagacagaagaagtaacttggccaccctgttgagaccagaag 120
 Db 165 ACTCATGAAGTACTTGAACAGCACAGAGGATCTTGGCCACCTGTGTGACCCAGCG 224
 Oy 121 cagtgacatacctcctcccgatgctcttgagcctatgacgctgacttctcttggtggggaat 180
 Db 225 CAGTGACCTATCCCTTCGCTGCTGTGAGCTATGCGCTGATCTTCTGTGGGGTAAAT 284
 Oy 181 gggcaactctcttgatgacagtgatgtgcagacatacgaacttgaagaacacacacaa 240
 Db 285 GGGCAATCTTCTGTGATGATGTGATGTCGACATCAACATTTGAAGACACCCACCA 344
 Oy 241 ctatcatctctcagcttggcagctccagatcgcgtgctcgtctcttgaggatgctct 300
 Db 345 CTACTATCTCTTACGCTTGGCATCTCAAGATCGCTGCTCTCTTGGGGATGCTCTT 404
 Oy 301 ggaatctcagagaatgagacaaattacaccttccctgctggccgtggagatgactact 360
 Db 405 GGAATCTCAAGATGAGGACAAATTACCTTCTCTGTGGGCTGTGGGATGCTACTT 464
 Oy 361 caagaagaccctctcgaagaactgtgctcttgcctccatctcagtgatccacagcttag 420
 Db 465 CAAGACAGCCCTCTTGAGAGCTGTGTGCTTGGCTTCATCTCAGTGTCAACAGGTAG 524
 Oy 421 cgttagagcgtatgctgagcactgttccaccccttcgagacgaagctggagagcagcgcg 480
 Db 525 CGTAGAGCGGTATGTGGCCCTTGTCCACCTTTCGAGCCAACTGGAGAGACACGCGCG 584
 Oy 481 agggagccctcagagatcctcagcctagcttgagagctctcgttgctcttcttgcacaa 540
 Db 585 ACGGGCCCTCAGGATCCTCAGCTAGCTGAGACCTTCTCTGTGTCTTCTTCCCA 644

Oy 541 taccagatccatgagatcaagttccagcacttcccacagggtccctcgtacctgctc 600
 Db 645 TACCAGATCATGATGATCAAGTTCCAGCACTTCCCAAGGGCTCCGTACTCGTGGCT 704
 Oy 601 agccacctgacagctcaaccacacacatggtggtatgaattatcatcatcaagctacag 660
 Db 705 AGCCACCTGCAGATCAACCAACCTATGGGTATTAACCTTATCATCACTCAAC 764
 Oy 661 ctccctcttaccatccctcccaatgaaccctcacaagcgttccctactactcattgagct 720
 Db 765 CTTCCTCTTCTACATCTCTCCCATGACCCCTCATCAGGCTCTTACTACTCATGGGGCT 824
 Oy 721 caggctgaagaagatgaatcccttgaggcgaacaaagtgcgtgataatccagacc 780
 Db 825 CAGGCTGAAGAAGATGAATCCCTTGAGCGCAACAAAGTGGCTGTAATTCACAGACC 884
 Oy 781 ctctagaagtcagatcccaagaatgctgttcttcttgcttcttctgttcttccctcgtg 840
 Db 885 CTCTAGAAAGTCAGTCAACCAAGATGCTGTGTGCTCTGTGCTGTGCTGCTGCTG 944
 Oy 841 gaccctctcagatgtagccgctctctcagcttctgagaaagtgaagacagatccct 900
 Db 945 GACCCCTTCATGATGAGCGGCTCTTCTTACCTTGTGGAAGAGTGACAGATCCCT 1004
 Oy 901 ggcctgctgttcaacctcaccatgtagatcagatgctctcttcttcttctgagctcgc 960
 Db 1005 GCGTGGCTGTTCACCTCATGATGATGATGATGATGATGATGATGATGATGATG 1064
 Oy 961 ggtcaaccccatatctatactcctcgtctgctgagccttccggggcgcttctgaaatg 1020
 Db 1065 GGTCAACCCCATATATATATACCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTG 1124
 Oy 1021 tgtctccctcactgcaaatggtgacatcccgagatccgacaggaactccagacca 1080
 Db 1125 TGCTCTCCCTCACTGCAAAATGGTGCCATCCCGGATCAGCCACAGGACTCCAGCCCA 1184
 Oy 1081 gaagatcatctcttgacagaatgtaacactcgttgagagctgacagaagaatgcaagcccca 1140
 Db 1185 GAAGATCATCTTCTGTGACGAATGTCACTCATGAGACTGACAGAGATGACAGGCCCA 1244
 Oy 1141 gtccctcgtgcatcatcatccacacacacacacacacacacacacacacacacacac 1200
 Db 1245 GTTCCCTGATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 1304
 Oy 1201 ggtacacata 1210
 Db 1305 GGTACCATGA 1314
 RESULT 2
 AF242875 1188 bp mRNA ROD 13-JUL-2000
 LOCUS Rattus norvegicus neuromedin U receptor 2 (NMU2R) mRNA, complete
 DEFINITION
 cds.
 ACCESSION AF242875
 VERSION AF242875.1 GI:9082157
 KEYWORDS
 SOURCE Norway rat.
 ORGANISM Rattus norvegicus
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 1188).
 REFERENCE
 AUTHORS Howard, A.D., Wang, R., Pong, S.S., Mellin, T.N., Strack, A., Guan, X.M.,
 Zeng, Z., Williams, D.L., Jr., Feiguer, S.D., Nunes, C.N., Murphy, B.,
 Stalf, J.N., Yu, H., Jiang, Q., Clements, M.K., Tan, C.P., McKee, K.K.,
 Hreniuk, D.L., McDonald, T.P., Lynch, K.R., Evans, J.F., Austin, C.P.,
 Caskey, C.T., Van der Ploeg, L.H. and Liu, Q.
 TITLE Identification of receptors for neuromedin U and its role in
 feeding
 JOURNAL Nature 406 (6791), 70-74 (2000)
 MEDLINE 20351041
 REFERENCE 2 (bases 1 to 1188)

AUTHORS Liu, Q., McDonald, T. P., Wang, R., Jiang, Q. and Howard, A.D.
TITLE Direct Submission
JOURNAL Submitted (09-MAR-2000) Pharmacology, Merck Research Labs, West
Point, PA 19486, USA
FEATURES Location/Qualifiers
source 1..1188
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
gene 1..1188
/gene="NMUR2"
CDS 1..1188
/gene="NMUR2"
/note="FM-4"
/codon_start=1
/product="neuromedin U receptor 2"
/protein_id="AA82736.1"
/db_xref="GI:9082158"
translation="MGKLENSAWIHDPMLKYNSTEXLAHLCGRKSDLSLPVSVA
ALIFVGMGNLVCNIVRHQTLPTNLYLFSLAVSDLLGLKSPLEIYEMHNY
PLRGPGVGYFTALFEVVCPSALISVTVEVVAIVHPRAKISERRARRRRLIS
LMSFSVFEFLPNTSINGIKRQHPFNSSVPSATCTYTKRMVYNLIOATSELYI
LPMFLISVLYILMGRLKRDSELEANKVAANIHRPSVTKMLFVLVLAICWTFP
HVDRLFSEVEEMTESLAAYFNLIHVVSGVEFLYSAVNPILNLSRFAARMNV
SPCKMCHPRHRPOGPPAQKIIIFLTECHLVLETDAGPQFPOSSIHNTNLTAPCAG
EVP"
BASE COUNT 243 a 361 c 278 g 306 t
ORIGIN
Query Match 96.4%; Score 1186.4; DB 10; Length 1188;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1187; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 23 atgggaaacttgaaatgctctcgatccagatcctctcaatgaagtaacttgaaacagc 82
Db 1 ATGGGAAACTTGAAATGCTCTCGATCCAGATCCACTGATGAAGTACTTGAACAGC 60
QY 83 aacagagagacttgcccaactgtgtgagacccaagcgacagtaactccttcgcgag 142
Db 61 ACAGAGAGACTTGCCCACTGTGTGACCCAAAGCGACGTACCTCCCTCCGGGG 120
QY 143 tctgtgacatacgctgacatctccctgtgtgaggaatagtgaaactcttcgtgtcat 202
Db 121 TCTGTGGGCTATGCGCTATCTCTCTGGGGGGTAATGCGAATCTTGTTGTCAG 180
QY 203 gtgattgtccgacatcagaacttgaagaaccccaactactatctcttcagcttgag 262
Db 181 GTGATTGTCCGACATCAGACTTTGAAGACACCACCACTACTATCTTTCAGCTTGG 240
QY 263 gtctcaagatctggtctctgtctctgtgggaatgctcttgaaatctcaagaatgtggc 322
Db 241 GTCTCAGATCTGCTCTGCTCTTGGGGGATGCTCTGGAATCTACAGATGTGGC 300
QY 323 aattacccttctcgttcgggacgtgtgagatctacttcaagaacagcccttcgagact 382
Db 301 AATTACCTTTCCTTGTGGGCTGTGGGATCTACTTCAACAGACCCCTCTTCAGACT 360
QY 383 gtgtgcttgcctcattctcagtgatccacaggttagcgttagagcgtatgttgccatt 442
Db 361 GTGTGCTTGTGCTCATCTCTCAGTCAACAGGTTAGGTTAGAGCGGTATGTGGCATT 420
QY 443 gtccacccttccgagccaagcttgagagacagcgagcgagccctcaagaatcctcagc 502
Db 421 GTCCACCTTTCGAGCAAGCTGAGAGACAGCGCGACCGCCCTCAGATCTTCAC 480
QY 503 ctgagcttgagactctcgtgtgtcttctcttgcaccaataacagatcagtgacatgaag 562
Db 481 CTAGCTGTGAGACTTCTGTGTGTCTTTCTTCCCAATACAGATCATGTGATCAAG 540
QY 563 ttcacgaacttcccaacggtctcctcgatcagctcagccacactgacatgacacaa 622
Db 541 TTCAGACACTTCCCAACGGGTCTCGTACCTGAGTCCAGCCACTGCACATCAACAA 600

QY 623 cccatgtgaggtgtataactgtatcatccaaagctaccagcttctcttctacatctccca 682
Db 601 CCCATGTGGGTATTAATCTGATCATCCAAAGTACCAAGCTTCTCTTCTACATCTCCCA 660
QY 683 atgaccctatcagctctcctactactacatcgtgaggtcgaagcttgaaagatgaatcc 742
Db 661 ATGACCCTATCAGGCTCTCTCTACTACATCCATCAGGGGCTCAGGCTGAAGAGATGAATCC 720
QY 743 ctgtgagcgaaacaaagtgtgtgtgaatatcacaagccctcagaagaatcagtaaccaag 802
Db 721 CTYTGAGCGCAACAAAGTGTGTAATATTCACAAACCTCTAGAAAGACATCAACAAAG 780
QY 803 atgctgtgtgtctgtgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 862
Db 781 ATGCTGT 840
QY 863 ctctctctcagcttctgtgaagaagtgaacagatccctcgtcgtgtgtgtgtgtgtgtgt 922
Db 841 CTCTCTCTCAGCTTGTGTGAAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
QY 923 catgtgtgtatcaggt 982
Db 901 CATGTGTGTATCAGGT 960
QY 983 ctccctgtctcggcgtctccggcgcgctcttgaaatgtgtgtgtgtgtgtgtgtgtgtgt 1042
Db 961 CTCTGTGTCTCGCGCTTCCGGCGCGCTTTCGAAATGTGTGTGTGTGTGTGTGTGTGTGT 1020
QY 1043 tgcacatcccgagcatcgcgcacacagaggaacctccagccacagaagaatcatctcttcagaga 1102
Db 1021 TGCATTCGCCGCGCATCGGCACACAGGAGACTCCAGCCACAAATATCATCTTGTGAACAGA 1080
QY 1103 tgtcaactctgtgtgagctgcagagagatgcaggccccagcttccctgtgtcatcattcc 1162
Db 1081 TGTCACTCTGTGAGACTGCACAGAGATGCAGGCCCCCACTTCCGTCAGTATCATCATC 1140
QY 1163 caaacacacaaacttaccacagccccctgtgagagagaagttaccataa 1210
Db 1141 CAAACACCAACTTACCAAGGCCCCCTGTGACAGAGAGATCACTAA 1188
RESULT 3
AF272363
LOCUS AF272363 1298 bp mRNA PRI 23-OCT-2000
DEFINITION Homo sapiens neuromedin U receptor 2 (NMUR2) mRNA, complete cds.
ACCESSION AF272363
VERSION AF272363.1 GI:10946202
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1298)
Raddatz, R., Wilson, A.E., Artymyshyn, R., Bonini, J.A., Borowsky, B.,
Boreju, L.W., Zhou, S., Kouranova, E.V., Nigorny, R., Guevarra, M.S.,
Dal, M., Lerman, G.S., Vayse, P.J., Branchek, T.A., Gerald, C.,
Porray, C. and Adam, N.
Identification and Characterization of Two Neuromedin U Receptors
Differentially Expressed in Peripheral Tissues and the Central
Nervous System
JOURNAL J. Biol. Chem. 275 (42), 32452-32459 (2000)
PUBMED 10899166
REFERENCE 2 (bases 1 to 1298)
Bonini, J.A., Raddatz, R., Wilson, A. and Borowsky, B.
Direct Submission
TITLE Submitted (25-MAY-2000) Target Discovery and Assessment, Synaptic
Pharmaceutical Corporation, 215 College Road, Paramus, NJ 07652,
USA
FEATURES Location/Qualifiers
source 1..1298
/organism="Homo sapiens"
/db_xref="taxon:9606"

gene
CDS
/chromosome="5"
/map="5q31.1-q31.3"
1..1298
/gene="NMUR2"
27..1274
/gene="NMUR2"
/note="G protein-coupled receptor"
/codon_start=1
/product="neuromedin U receptor 2"
/protein_id="AAG24794.1"
/db_xref="GI:10946203"
/translation="MSGMKLQNASMIYQOKLEDPQKHINSTEYLAFLCGRSHF
FLPSVAVPIFVGVIGNVLCLVILQHQAKTPNLYLFLSLAVSDLVLLGLP
VYEMRNYPFLFPGVCEKTALEFVCFASLITTVSVERVVAIILHFRKLOSTR
RAALRIILGWGFSVLESLPNTSINGIKFHYHPNSLPGSATCIVIKPMIYNFIQ
VTSFLYFLPMTVIVLYLMLRLKDKSLDAGNANIORPKRSVKMKFLVAV
FALCAPPHIDILFSPVEMESLSLAAPNIVHYVSGVFYSSVNPRIYLLSRF
QAFONVSSFRKHSHSDPOLPPAQRNIFLIECHFEVLIEDIDGPPFCOSSMHNH
LPTALSSQMSSTNYQSFHNKT"
BASE COUNT 278 a 405 c 282 g 333 t
ORIGIN

Query Match 65.0%; Score 799.6; DB 9; Length 1298;
Best Local Similarity 79.2%; Pred. No. 8e-211;
Matches 968; Conservative 0; Mismatches 239; Indels 15; Gaps 1;

QY 5 tggatttaagctcagtagtggaagaacttgaaatgcttcctcgatccac----- 55
DB 18 TGGATTTTTAAATGTCAGGAGATGCAAAACCTTCAGAAATGCTTCGAGATCTACAGCAGCAAAA 77
QY 56 -----gatacctctca tgaagtaacttgaacaagacagagagtagtctggccacactgtc 109
DB 78 CTAGAAGATCATTCAGAAACACCTGGAACAGCAGCAGAGAGATCTGGCCTTCCTCTCG 137
QY 110 ggaaccaagcagtagtaccatccctccggtgtctgtgctgacctagcgctgactctcc 169
DB 138 GGACCTCGGCGGCGACCACTTCTCCCTCCCGGTCTGTGTATGTGCAATTTTGTG 197
QY 170 gtaggggtaagtggaactctctgtgtgtcagtagtattgtcgaataagacttgaag 229
DB 198 GTGGGGGTATTTGGCAATGCTCTGTGTGTCTGTGTATTTGACAGCAGCGGTATGAG 257
QY 230 aacaccacaactatctctctcagacttgcagctcagatctcgtgtccctctg 289
DB 258 AGGCCACCAACTACTACCTTCTCAGCCTGGCGGTCTGACTCTCGTGGTCTGCTCT 317
QY 290 gggagtctctggaaatctaagagatgtggcacaattacccttccctgtcggacctgtg 349
DB 318 GGAATGCCCTGGAGGCTCTATGATGTGGCGCACTACCTCTTCTTGTTCGGGCCCTG 377
QY 350 ggaagtactcagaacagccctctcgaagactgtgtcttgcctcccaacttccagtg 409
DB 378 GCGTGTACTTCAAGACGGCCCTCTTTGAGACCGGTGTCTGCTCCCTCTCATCTCAGACATC 437
QY 410 accaaggttagcgttagagcgctatgtgcacattgtccaccccttccagagccaagctgag 469
DB 438 ACCACGCTCAGCGTGAAGCGCTACGTGGCCATCTACACCCGTTCCGGCCAAATGAC 497
QY 470 aggaagcggcgaagggccctcagatccctcagcagctagcttgtagctctctgtgtcctt 529
DB 498 AGCACCCTGGCGGCGGCTCTCAGGATCTCGGCGATGCTGTGGGGCTTCTCCGTCCTTC 557
QY 530 tcttgcacaataacagatccaatgacatcaagtttcagcagcttcccaagggctctcc 589
DB 558 TCCCTGTCCCAACACGACATCCATGGCATCACTTCCACTACTTCCCAATGGGCTCCG 617
QY 590 gtaactgtgctcagcaccctcgtcacaagtcacaaacccaatgtgggtgatalaactgtatcac 649
DB 618 GTCCGAGGTTCCGGCAGCTGTAGGTCATCAACCCATGTGATGATCAATTTCAATTCATC 677
QY 650 caagctacagctctctctcttacaatccctcccaatgaacctcaagcgctcctactaac 709

DB 678 CAGGTCACTCTCTTCTTATTCACCTTCCTCCCATGATGATCATGATGTCTCTACTAC 737
QY 710 ctcatgggctcagagctgaagagatgaatcccttgaagcgaaacaaatgtgctgaat 769
DB 738 CTCATGGCACTGACACTAAAGAAAGCAATCTCTTGAGGACAGATGAAGGATGCCAAT 797
QY 770 attcacaagccctctgaagaagtcacacccaagaatgtctgttcttggctcctgtgtt 829
DB 798 ATTCAAAGACCTCGCAAGAAATCAGTCAACAGATGCTGTTGTCTGGCTAGTGTTT 857
QY 830 gccatctgtgcagcccccctccatctgtgacggcgctcttctcagcttgtgaaagatg 889
DB 858 GCATCTGTGGGCCCCCGTTCACATGACCGACTCTTCTTCAGCTTGTGGAGGAGTG 917
QY 890 acaagatccctgtcgtcgtgtcacaaccatccatcagtgatcagagttctcttctat 949
DB 918 AGGAATCCCTGGCTGTGTGTCAACCTGTCATGTGTGTGTGTGTGTGTGTGTGTGT 977
QY 950 ctgagctcggcggttcaaccatatactataaactcctgtctcgcgctcgcggcgcc 1009
DB 978 CTGAGCTCAGCTGTCAACCCCATTAATCTATACCTACTGTCTGCGCGCTTCAGGACGA 1037
QY 1010 ttctgaaatgtgtctccctaccctcagatgtgtccatcccgcatcgcccaagga 1069
DB 1038 TTCCAGAAATGTGATCTTCTTCTTCCACAAACAGTGGCACCTCCAGACCAAGATTG 1097
QY 1070 cctccagcccaagagatcatctcttctgaagaaatgtcacctgtgagctgagctgaagagat 1129
DB 1098 CCACCTGCCCCAGCGGAACATCTTCTCTGACAGAAATGCCACTTTTGAAGCTGACCGAAGAT 1157
QY 1130 gcagcccccagcttccctgtgtcagtcacatccatccacaacccaacttccacggccccc 1189
DB 1158 ATGAGTCCCAATTCATTCATGATGATGATGATGATGATGATGATGATGATGATGAT 1217
QY 1190 tgtgcagagaggtaccataa 1211
DB 1218 TCTAGTGAACAGATGTCAAGAA 1239

RESULT 4
AX109234 LOCUS 1658 bp DNA PAT 30-APR-2001
DEFINITION Sequence 1 from Patent WO0125269.
ACCESSION AX109234
VERSION AX109234.1 GI:13924107
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1658)
AUTHORS Delezeny, M., Berger, C., Loeken, C., Nys, G. and Venema, J.
TITLE Human g-protein coupled receptor
JOURNAL Patent: WO 0125269-A 1 12-APR-2001;
Solvay Pharmaceuticals B.V. (NL)
LOCATION/Qualifiers
1..1658
/organism="Homo sapiens"
/db_xref="taxon:9606"
55..1302
/note="IGS4A long version"
/codon_start=1
/protein_id="CAC37771.1"
/db_xref="GI:13924108"
/translation="MSGMKLQNASMIYQOKLEDPQKHINSTEYLAFLCGRSHF
FLPSVAVPIFVGVIGNVLCLVILQHQAKTPNLYLFLSLAVSDLVLLGLP
VYEMRNYPFLFPGVCEKTALEFVCFASLITTVSVERVVAIILHFRKLOSTR
RAALRIILGWGFSVLESLPNTSINGIKFHYHPNSLPGSATCIVIKPMIYNFIQ
VTSFLYFLPMTVIVLYLMLRLKDKSLDAGNANIORPKRSVKMKFLVAV
FALCAPPHIDILFSPVEMESLSLAAPNIVHYVSGVFYSSVNPRIYLLSRF
QAFONVSSFRKHSHSDPOLPPAQRNIFLIECHFEVLIEDIDGPPFCOSSMHNH
LPTALSSQMSSTNYQSFHNKT"

BASE COUNT 377 a 490 c 346 g 445 t

Db	226	GTGGGGGTCATTTGGCCAAATGTCTCTGTGTGCTGTGTATTTCTGCAGCACACCGCTATTGAAAG	285
QY	230	aaaccccaaccaatactatctcttcacgctcttgagactctgagactcagaatctgctgctcgtctcttg	289
Db	286	ACGCCACACCACTACTACTACTCTTCAGCGCTGGCGGCTCTCTGACGCCCTGGTCTGCTCTAC	345
QY	290	ggagatgctctctcggaatcttaagaaatgtgtgcaacaataacccttcctcgtctcgagctgtg	349
Db	346	GGAAATGCCCTTGAGAGTCTATGAATATGTGGGGCAACATACCTTTCTTGTTCCGGGGCCCGTG	405
QY	350	ggaatgactactcaagaagaagccctcttcgagaatgtgtgctcttgccctcatcttcacgtatc	409
Db	406	GGCTGCTACTTCTTAAGACGCGCCCTTTTGAAGACGCTGTGCTTCCGCTCTCATCTCTCAGCATC	465
QY	410	accacagcttaagcgttaagagcgtatgtgtgacatctgttccaacctcttcgagacaaagcttggag	469
Db	466	ACCAACCGTCACACGTTGAGAGCGCTACTGTGGCCATCTCATACCCCGTTCCGGCCAAATGACAG	525
QY	470	agcaacgagcgaaagcgggcccctaaagatccctcagccttagctctgagactctctgtgtgctctt	529
Db	536	AGCAACCGCGCCCGGGCCCTTAGGATCTCTGGGCAATCTGTGGGGCTTCTCCGTGCTCTT	585
QY	530	tctttgcccacaatcacagacatccatcgagatcaagttccagacacttcccacacgggtctctcc	589
Db	586	TTCCTGCCCCAACACACGACATTCATGGCATCAAGTTCCACTACTTCCCCCAATGGGTCCCTG	645
QY	590	gtacctgtgctcagccactctgacacatccacaacaaaccaaactgtgtggtgtataacttgatcatc	649
Db	646	GTTCCAGGTTGGGCGACACTGTACGTCATCAAGCCCATGTGGATCTCTCAATTTTCATCATC	705
QY	650	caagctaacagactctcctctcttaatactctcccaatgtgacctcaatcagcgtcctcctactac	709
Db	706	CAGGTCACTCTCTTCTTAATTTACTCTCTCCCAATGAGCTGTCATCATGTCCTCTCTACTAC	765

LOCUS	AX109238				
DEFINITION	Sequence	5 from Patent WO0125269.			
ACCESSION	AX109238				
VERSION	AX109238.1	GI:13924111			
KEYWORDS					
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Enkharjola; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1 (bases 1 to 1658)				
TITLE	Deleensnllder, W., Berger, C., Loeken, C., Nys, G. and Venema, J.				
JOURNAL	Human g-protein coupled receptor				
	Patent: WO 0125269-A 5 12-APR-2001;				
	Solvay Pharmaceuticals B.V. (NL)				
FEATURES	Location/Qualifiers				
source	1..1658				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	55..1302				
	/note="IGS4B long version"				
	/codon_start=1				
	/protein_id="CAC37773.1"				
	/db_xref="GI:13924112"				
	/translation="MSGSEKRLQNSWYIOOKLEDPFOKHLNSTFEYLAFICGPRRSIF				
	FLPVSVPVPLEVGVGVGNVLVCYLIIQHQMKRPNTVYSLVASDLVLILMPLE				
	VYEMRNVPFLGEGVGVGFKTALFETVCFSALISITVYSERYVALIHPFAKIQSTR				
	RRARILGIVMGFSVLFSLPNTSIHGKIFHFHPGSLVPGSANTCVIKPMVINFIIQ				
	VTSLEFLIPMTVLSVLYLALRLKDKSLSEADEGNANIQRPCRSVNNKLEFLVLYLV				
	FAIQMAQPHIDIRLFFPSVEEWETSELAAVENLVHVSGLVFLSSAVNPDIYNLSRFR				
	QAACQNTYSFPHKQMHSDHQDQLPQAQRNIFLETCHFEVLLEDIGPQFLCQSSVHNSH				
	LPTNLSSEQMKRTYIOSEHFNKT"				

QY 470 agcagcgcgacgagccctcagatctccagctacgtctctctgtgctctt 529

[illegible]

LOCUS	1248 bp	PRI	21-SEP-2000
DEFINITION	AB041228 Homo sapiens mRNA for G protein-coupled receptor TGR-1, complete cds.		
ACCESSION	AB041228		
VERSION	AB041228.1 GI:10257380		
KEYWORDS	G protein-coupled receptor TGR-1.		
SOURCE	Homo sapiens cDNA to mRNA.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
AUTHORS	1 (sites) Hosoya,M., Moriya,T., Kawamata,Y., Ohkubo,S., Fujii,R., Matsui,H., Shintani,Y., Fukusumi,S., Habeta,Y., Hinuma,S., Onda,H., Nishimura,O. and Fujino,M.		
TITLE	Identification and Functional Characterization of a Novel Subtype of Neutrokinin U Receptor		
JOURNAL	J. Biol. Chem. 275 (38), 29528-29532 (2000)		
PUBMED	10887190		
REFERENCE	2 (bases 1 to 1248)		
AUTHORS	Shintani,Y., Moriya,T., Ohkubo,S. and Matsui,H.		
TITLE	Direct Submission		
JOURNAL	Submitted (03-APR-2000) to the DDBJ/EMBL/GenBank databases. Yasushi Shintani, Yakeda Chemical Industries, LTD, Discovery Research Laboratories 1: 10 Wadai, Tsukuba, Ibaraki 300-4293, Japan (E-mail:Shintani.Yasushi@yakeda.co.jp, Tel.81-298-64-5011, Fax:81-298-64-5000)		
FEATURES	Location/Qualifiers		
source	1..1248		
gene	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	10..1248		
CDS	/gene="TGR-1"		
	10..1248		
	/gene="TGR-1"		
	/codon_start=1		
	/product="G protein-coupled receptor TGR-1"		
	/protein_id="BA013721.1"		
	/db_xref="GI:10257381"		
	/translation="MEKLNQNSWIVYQKLEDPQKHNSSTEEYLAFICGPRSHFFLP		

	Query Match	64.2%;	Score 790.8;	DB 9;	Length 1248;	
	Best Local Similarity	79.2%;	Pred. No. 2.2e-208;			
	Matches 958;	Conservative	0;	Mismatches 237;	Indels 15;	Gaps 1;
BASE COUNT	269 a	393 c	269 g	317 t		
ORIGIN						
OY	17	tcaagtaatgaggaaacttgaaaabctcttccttgcataccac-----gatcct	61			
Dd	4	TCAGGGAATGGAAAACTTCAGAAATCCTTCCGTGACATCTACAGACAGAATACTGAAGATCA	63			
OY	62	cctcatgaagtacttgyaacagcacagagaagiacttggcccacctgttlytggaocecaagcgc	121			
Dd	64	TTCAGAGAACACTTGAMAGCACCGGAGGAGATATCGGCCTTCCTCGCGACCTCGCGCG	123			
OY	122	agtgcaccatcccttcgcggtgtctgttggccatatgcctgcattcttccttggtygggtaatg	181			
Dd	124	AGCCACTTCCTCCCTCCCGTGTCTGTGGGTGATGTGCCAATTMTTGTGTGGTGGGCTCATTT	183			
OY	182	ggcaatctctctgtgtgcatagttatcttgcacatacgaactttgaaagacaaccaaac	241			
Dd	184	GGCAATGTCTCTGTGTGCTTGGCTGTGTATTTCTGCAGCACCGGCTATGAACAACGCCACCAC	243			
OY	242	tactatctcttcagcttggcagtlctcagaatctgctgtgtccttcttlttggtagtgccttg	301			
Dd	244	TACTACTCTTTCAGCGCTGCGGCTGTCTGTACCGTCTGTGTCGTCGCTCTTGGATGGCCCCTG	303			
OY	302	gaaatctcagagatgttggcacaataacccttctcctgttcgggctcttggatgtctacttc	361			
Dd	304	GAGGCTCTATGAGANTGTGGCGCAACTACCCCTTCTTGTGTTGGGGCCCGTGCGTCACTTC	363			
OY	362	aagaagcccccttcctcgagactgttgccttcttcctcatcttcacatgctacacagcttgc	421			
Dd	364	AAGAGCGCCCTCTTTGAGACCGCTGTGTGTGCTTGCCTTCATCTTCAGCATACCAACCGCTCAGC	423			
OY	422	gtlagaagctatgttggccatttgtccaaccttctcgagcacaagcttgyaagacagcgcgga	481			
Dd	424	GTGGAGCGCTAGTGGCCATCTTACACCCGTTCCCGGCCAAMAATGCAGAGCACCGCGCGC	483			
OY	482	cgggccccctcagatctcctaagccttagtcttggagctctctctgttgccttcttgcacat	541			
Dd	484	CGGGCCCTCAGGATCTTGGCATCTCTGCGGGGCTTCTCCGTCTTCTCCCTCGCCCAAC	543			
OY	542	accagatcataitgycatalcaagtcttcagagacttcccaacgggtctccgtaccctgcgtcca	601			
Dd	544	ACCAAGATTCATGGCATGATCAAGTCCACACTTCTCCCAATGGGTGCTGCTCCAGTTTCG	603			
OY	602	gccacctgcacagtlcaccaacaaccactgttgggttgtaataactgatalcatcaagcttaccagc	661			
Dd	604	GCCACTCTTAGCGGTATCATAGGCCATGTGGATGCTTAACAATTCATCATCATCAGGTACCTCC	663			
OY	662	ttcctctctacatctctcccaatgacccctcaatcagcgtctcctctactaccatcaygggctc	721			
Dd	664	TTCCTATTCTACACCTCTCCCATGACTGTCTCATGACTGTCTCTACTACTCATGTGGACTTC	723			
OY	722	aggcctgaagagagatgaatctctttagagcgaaacaagaggcttgygaatatcaagaagacc	781			
Dd	724	AGACTTAAAGAAAAGCAAAATCTCTTGAGGCAATGAGAAAGGAATGCAAAATATTCAAAGACC	783			
OY	782	ctctagaagctcaagtlcaccaagaatgctgtcttcttgcctccgtgttgcatactgcctcgcg	841			
Dd	784	TGCAGAAATATCATGTAACAAGATGCTGTTGTGTCTTGCGCTTAAGTCTTTCGATCTGTTGG	843			
OY	842	aacctctcatgttgcacagcgctcttcttcagcttlttgyaagagtgagacagatcccttg	901			

Query Match	64.0%	Score 788	DB 9	Length 1239
Best Local Similarity	79.2%	Pred. No. 1.3e-207		
Matches 954	Conservative 0	Mismatches 235	Indels 15	Gaps 1

[illegible]

||||| ||||||| ||| ||||||| || ||| ||| ||||| |||||
Db 1081 ATCTCCGACAGATGACCTTTGTGGAGCTGACCGAAGATATAGTCCCAATTCCCA 1140
Oy 1148 ggtcagtcattccatccacacacacacttaccagccctctgacggagaggtacc 1207
Db 1141 TGTCAGTATCATGACACACTCTCACCACAGCCCTCTTGTGTAACAGATGTCA 1200
Oy 1208 taac 1211
Db 1201 AGAA 1204
RESULT 11
AX109242 1594 bp DNA PAT 30-Apr-2001
LOCUS AX109242 Sequence 9 from Patent WO0125269.
ACCESSION AX109242
VERSION AX109242.1 GI:13924115
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Deleersnijder, W., Berger, C., Loeken, C., Ny, G. and Vanema, J.
TITLE Human g-protein coupled receptor
JOURNAL Patent: WO 0125269-A 9 12-Apr-2001;
Solvay Pharmaceuticals B.V. (NL)
FEATURES
Source Location/Qualifiers
1..1594
/organism="Homo sapiens"
/db_xref="taxon:9606"
55..945
/note="IGS4A truncated DNA long version"
/codon_start=1
/protein_id="CAC37775.1"
/db_xref="GI:13924116"
/translation="MSGEMKLNASMTYQKLEDPFOKHLNSTEYLAFLOGPRSHF
FLPVSYYVPIFVVGIVGNLVCLVLIHOHOMKPTNYLESLAVSDLVLLGMPLE
VYEMRNYPPLFGPGVCGKFKTALPEYGFASLITTSVSRYVAIIHPRAKQSTR
RRALRLIGIVMGSVLFSLEPNTSHGKIFHFPGLSVPSGATCTVAKPMITVFIIG
VTSFLFLPMTVISLVLYLMLALRLKDKSLEDEGNANIORPCRSKSVNKLSTMRSG
VNPWLICSTSMCOVSSST"
BASE COUNT 370 a 473 c 332 g 419 t
ORIGIN

Query Match 55.0%; Score 677.6; DB 6; Length 1594;
Best Local Similarity 74.8%; Pred. No. 5.7e-177;
Matches 914; Conservative 0; Mismatches 229; Indels 79; Gaps 2;

Db 346 GGAATGCCCTGAGTCTATAGATGTGGCGCAACTACCTTTCTTTGTTGGGCCCTG 405
Oy 350 gtagtactatcaagaagaccctcttgagaactgtgtgtctgtccatctcagtg 409
Db 406 GGCTGCTACTTTCAGACGGCCCTCTTTGAGACCGTGTCTTCCCTCATCTCAGATC 465
Oy 410 accaagtgtaagcgtatagcgtatgtgacatltgtccacacttctcgagcaagctgag 469
Db 466 AOCACGTCACGCTGAGACGGCTACTGCGCATTCCTACACCGGTTCCGGCCAACTGAG 525
Oy 470 agcacgagcgacgagccctcagaagatcctcagcctagtcgtgagactctctgtgtct 529
Db 526 AGCACCGCGCGCGCCCTCAGGATCCTCGGCATCTGTGGGCTTCTCCGTCTTTC 585
Oy 530 tcttgcccaataccagatccatgagcctgaagttccagcacttcccaagggtcccc 589
Db 586 TCCCTGCCCAACACGAGATCCATGGCATCAAGTTCACACTTCCCAATGGGCTCTG 645
Oy 590 gtaccctgctcagccacgtcgaagtcacacaaaccatgtgtgtatcaactgtac 649
Db 646 GTCCAGGTTGCGGCACCTGTACGTCATCAAGCCCATGTGGATCTACATTTATCATC 705
Oy 650 caagctaccagcttctcttctacatcctcccaatgaccctcatcagcgtcctctac 709
Db 706 CAGGTACACTCTCTTCTTCTTCTACCTCTCCCATGATGATCATCATGATCTCTACTAC 765
Oy 710 ctcatggggctcagcgttaagagagatgaatccctctgagggcgaaaggtgctgtgat 769
Db 766 CTCATGGACACTGACACTAAGAACACAATCTCTTGGAGCGAGATGAAGGAAATCAAT 825
Oy 770 attcacagaccctcagaagtcagtcacacaaagatgtgtctgtctgtctgtct 829
Db 826 ATTCAAAACCCCTGAGAAATCACTCAACAAAGATGCTG----- 865
Oy 830 gccatctgtgagacccctccatgtgagccgctctctcagcttctgtgagagtg 889
Db 866 -----CTTGTGTGAGAGATG 881
Oy 890 acagagtcctcgtgctgtgttcaactcctacatgtgtatagtggtctctcttat 949
Db 892 AGTGATCCCTGGCTGTGTCTTCAACTCGTCCATGTGTGAGTGTCTTCTTCTAC 941
Oy 950 ctgaagctcgcgttcaaccccatctacataacccctcgtctcgcgtccgcgcgcgc 1009
Db 942 CTGAGCTCAGCTGTCAACCCCATTTATTAACCTACTGTCTCGCCGCTTCCAGGAC 1001
Oy 1010 ttctgaatgtgtctccctacactgcaaatgtgtccatcccgagcatcgccacagga 1069
Db 1002 TTCCAGATGTGATCTCTTTCCACAAACAGTGCGACTCCGACATGACCCACAGTTG 1061
Oy 1070 cctcagcccgagaagatcatcttcttgacagaaatgtcaactcgtgtgagctgaagagat 1129
Db 1062 CCACCTGCCACACGGAACATCTCTGTACAGAAATGCCACTTGTGGAGCTGACCCAAAT 1121
Oy 1130 gcaagcccgagtcctcgtgtgagtcacatccacacacacacacacacacacacac 1189
Db 1122 ATAGTCTCCCAATTTCCATGTACATCATCATGACCAACTCTCAGCTCCCAACGCTC 1181
Oy 1190 tgtcagagaggtacataa 1211
Db 1182 TCTAGTGAACAGATGTCAAGAA 1203
RESULT 12
AX109244 1594 bp DNA PAT 30-Apr-2001
LOCUS AX109244 Sequence 11 from Patent WO0125269.
ACCESSION AX109244
VERSION AX109244.1 GI:13924117
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 1594)
 AUTHORS Delaunay, W., Berger, C., Loeken, C., Nys, G. and Venema, J.
 TITLE Human g-protein coupled receptor
 JOURNAL Patent: WO 0125265-A 11 12-APR-2001;
 Solvay Pharmaceuticals B.V. (NL)
 FEATURES
 source Location/Qualifiers
 1..1594
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 64..945
 CDS
 /note="IGS4A truncated DNA short version"
 /codon_start=1
 /protein_id="CAC37776.1"
 /db_xref="GI:13924118"
 /translation="MEKIQNMSWITQKLEDPFOKHLNTEELAFICGPRSHPELP
 VSVVYPIFVNGVIGNVLYCVLIQNAKTPNTLYESLAYSDLVLLGMPLEYE
 MRNPFLGPGVGTCTALFETVCPASLITVSEVRYVAILHFRKALQSTERRA
 LRILIGWGSVLSLPNTSINGIRFHPNGSLVGSATVIRKPMIYNFIQVTS
 FLFLYLPMTVISLVLYLMALRLKDKSLSEADGNANIORPKRSVKNMLSLWRSQVNP
 MLCTSSMNCVSSST"

BASE COUNT 370 a 473 c 332 g 419 t
 ORIGIN

Query Match 55.0%; Score 677.6; DB 6; Length 1594;
 Best Local Similarity 74.8%; Pred. No. 5,7e-177;
 Matches 914; Conservative 0; Mismatches 229; Indels 79; Gaps 2;

OY 5 tggatttaagtcagtaatgggaaacttgaaatgcttcgtatgcac----- 55
 DB 46 TGGATTTTATGTGCGAGGAGGAAAACTTCAGAAATGCTTCGAGACTGACGAGAAA 105
 OY 56 -----gattccctcatgaagtactgaacagcagagagagacttgagccacctgtg 109
 DB 106 CTGAGAAGATTCATTCAGAAACACCTGACAGACGAGAGATCTGGCTTCCTGTC 165
 OY 110 ggaacgaagcagtgagactatccctccggtgctggtgagcctatgctgacttcc 169
 DB 166 GGACCTGCGGCGAGCCACTTCTCCGCCGTGTGTGTGATGAGCAATTTTGTG 225
 OY 170 gtgggggtaatgggaactctctgtgtgtaagtgatgtgctcgaaatcgaacttgaag 229
 DB 226 GTGGGGGTCTTGTGCAATGTGCTGTGTGCTGTGAGCAACCGGATGAG 285
 OY 230 aac 289
 DB 286 AGCCCCACCACTACTACTCTTACGCTGCGGTCTCTGACTCTGCTGCTGCTCT 345
 OY 290 gggatgctcttgaaactcgaagatgtggacaatbaacccttctcgttcgggctgtg 349
 DB 346 GGAATGCCCTTGGAGGTCTGTGAGATGTGGCAACCTTCTGTCGGGCCGTG 405
 OY 350 ggaatgctcttgaaagcagcctctcgaagacgtgtgtgtgtgtgtgtgtgtgtgt 409
 DB 406 GGTGTGCTACTTCAAGAGCGGCTCTTGTGAGACCGGTGTGCTTCCCACTTCCAC 465
 OY 410 aacacagtgtagcgtagagcgtatgtggccatgttccacacacacacacacacacac 469
 DB 466 ACACCGTCAAGCTTGAAGCGCTTGTGAGACCGGTGTGCTTCCCACTTCCAC 525
 OY 470 agcagcggcagcggcggcctcgaagatcctcgaagcgtgtgtgtgtgtgtgtgtgt 529
 DB 526 AGCAGCGGCGCGGCGGCTTCAAGATCTTGGGATCGTGTGGGCGTCTTCGTCGTC 585
 OY 530 tcttggcccaatcgaacatcgaatgcatcaagtctcgaacacacacacacacacac 589
 DB 586 TCCCTGCCCAACACGACATCCATGGCATCAAGTTCCACTTCCCAATGGGCTCTG 645
 OY 590 gtaccctggctcagcaccctcagacgtcacaacacacacacacacacacacacacac 649
 DB 646 GTCCCAAGTTCCGCGACCTGTACGCTCATCAAGCCCATGTGATCTACAAATTTCA 705

OY 650 caagctaccagcttccctctctacatccctcccaatgacccctcagctccctactac 709
 DB 706 CAGGTACACCTCTTCCATCTACCTCTCTCCCATGACTGACTGACTGACTGACTGACT 765
 OY 710 ctcatggggtcagcgtcgaagagatgaaatcccttgagcggaacacacacacacacac 769
 DB 766 CTCATGGCACTCAGACTAAAGAAAGACAAATCTCTTGAGCGAGATGAGGAAATGCA 825
 OY 770 attcagaccctcctcgaagagatgaaatgaaatgaaatgaaatgaaatgaaatgaa 829
 DB 826 ATTCAAGACCTTGGCAAAATCAGTCAACAGATGCTG----- 865
 OY 830 gccatctgctggaacccctcctcctcctcctcctcctcctcctcctcctcctcct 869
 DB 866 -----CTTGTGAGAGGTG 881
 OY 890 acagagtcctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 949
 DB 882 AGTGAATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 941
 OY 950 ctgagctcggcggctcaacccatctatcctaaactcctgctgctgctgctgctgctg 1009
 DB 942 CTGAGCTCAGCTTCAACCCCATTTATCTATACCTTCTGCTGCTGCTGCTGCTGCT 1001
 OY 1010 ttctgaaatgtgtctccctacacctgcaaatggtgcatccctcggatcggcagcag 1069
 DB 1002 TTCCAGAAATGTATCTCTTCTTCCACAAACATGTGGACATCCAGCATGACCAAG 1061
 OY 1070 cctccagcccaagaatcatctctctgacagaaatgtgtaactctgtgagctgacag 1129
 DB 1062 CCACCTGCCAGCGGAGAACATCTTCTGACAGAAATGCTGCTGAGCTGAGCAGCA 1121
 OY 1130 gcagggcccccagcttccctgtgctgctgctgctgctgctgctgctgctgctgct 1189
 DB 1122 ATAGTGTCCCAATTTCCCATGTCACTCATCATGACACACTCTGACTCCCAACAG 1181
 OY 1190 tgtcagagagagttaccataa 1211
 DB 1182 TCTAGTGAACAGATGTCAAGAA 1203

RESULT 13
 AC008571/c
 LOCUS Homo sapiens chromosome 5 clone CTC-550M4, WORKING DRAFT SEQUENCE,
 DEFINITION 4 unordered pieces.
 ACCESSION AC008571
 VERSION AC008571.4 GI:13699386
 KEYWORDS HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 215088)
 AUTHORS DOE Joint Genome Institute.
 TITLE Sequencing of Human Chromosome 5
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 215088)
 AUTHORS DOE Joint Genome Institute.
 TITLE Direct Submission
 JOURNAL Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
 Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94596, USA
 COMMENT On Apr 20, 2001 this sequence version replaced gi:7708957.
 -----Genome Center
 Center: Joint Genome Institute
 Center Code: JGI
 Web site: http://www.jgi.doe.gov

 Project Information
 Center Project Name: 396672, H361
 Center clone name: CIT-HSPC_550M4

 Summary Statistics

```

Consensus quality: 204117 bases at least Q40
Consensus quality: 210928 bases at least Q30
Consensus quality: 212668 bases at least Q20
Estimated insert size: 223000; pulse-field gel estimation
Estimated insert size: 224780; sum-of-contigs estimation
Quality coverage: 6.63 in Q20 bases; pulse-field gel estimation
Quality coverage: 6.89 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
20386 20385: contig of 20385 bp in length
20386 20485: gap of unknown length
20486 35143: contig of 14658 bp in length
35144 35243: gap of unknown length
35244 84715: contig of 49471 bp in length
84715 84814: gap of unknown length
84815 215088: contig of 130274 bp in length.
Location/Qualifiers

```

```

source
1. .215088
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /chromosome="5"
  /clone="C7C-550M4"
  /clone_11b="Caltech human BAC library C"
  /clone_41474_c_42096_g_64807_t_301_others
BASE COUNT      66410
ORIGIN

```

Query Match	39.1%	Score 481.4	DB 2:	Length 215088
Best Local Similarity	79.6%	Pred. No. 1.9e-122		
Matches 588	Conservative	0	Mismatches 136	Indels 15
				Gaps 1
QY 5	tgagatttaagctcagtaatggaaacacttgaaatgctccttgatccac-----	55		
Dh 105570	TGGATTTTAAATGTCAGGGATGGAATAAAGCTTCAGAAATGCTTCTGGATCTACCAAGCAGAA	105511		
QY 56	-----galtcctcaltgaagtaacttgacaagcacagaaagtaacttggccccactgtgt	109		
Dh 105510	CTAGAGAGATCCATTCTCAGAAACACCTGAAACACAGACAGAGATGATCTGACCTTCCTC	105451		
QY 110	ggaccacaagcggaagtgacactatccctccggtgtctgttgacctatgactgtctcgt	169		
Dh 105450	GGACCTCGGGCGGACGACACTTCTTCTCCCGGTGTGTGTGTGTATGTGCAATTTTGTGG	105391		
QY 170	gtvggggtaatbgygaatctctctgtgtgtgcatggtgtaatgtctcgaatcagaacttgaag	229		
Dh 105390	GTGGGGGTCATTGGCAATGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	105331		
QY 230	acaaccacaactactatctctctcgaacttgcagcttgcagatctcagatctgactgtctgt	289		
Dh 105330	ACGCCACCACTACTACTCTTTCAGCCGTGGGGTCTGTGACCTCTGCTGCTCTCTCTCT	105271		
QY 290	gggaagctcttgaaatactcaagagatbgygacaaatlaaccttctctgttctgggcctgtg	349		
Dh 105270	GGAAATGCCCTTGGAAGTCTATAGATGTGGCCCAACTACCTTTCTGTGTGGGCCCGTG	105211		
QY 350	ggaatctactactcaagaacagccctctctcgaagtgtgtgctcttgacctcaactcaagtgc	409		
Dh 105210	GGCTCTACTTCAACACGGCCCTCTTTGAGACCGGTGTGTGTGTGTGTGTGTGTGTGTGTGT	105151		
QY 410	accacggtlaagcgtatagagcctaibgygccaatgttccaccccttccagaccaagctbga	469		
Dh 105150	ACCACCGTCAGCGTGAAGGCGTACGTGGCCATCTCAACACCGTTCGCGGCCCAACGTGACG	105091		
QY 470	agcagaagcgagagggccctcagaactcgaacactagatctgaagctctctgtgtctt	529		
Dh 105090	AGCACCGGCGCGGGCCCTTCAGATCTCTCGGCATCTGTGTGGGGCTTTCACGTGCTCTTC	105031		

OY	530	tccttgcgccaataaccagatgccatcgaagatccaagtctccagcaactctcccaacggtctctcc	569
Db	105030	tccttgcgccaataaccagatgccatcgaagatccaagtctccagcaactctcccaacggtctctcc	1049711
OY	590	gtacctgtgtcagcaccactgtcacagtcacccaacccaatgtggtgtataactgcatc	649
Db	104970	gtccacagtttggccacctgtacggtctatcaggccattgttgatcttgcatttcattcattc	1049111
OY	650	caagctacagactctctctctctacatccctcccaatgacccctacacagcgtctctactac	709
Db	104910	cagctacacct	1048511
OY	710	ctctacgggtctcagcgtcga	728
Db	104850	ctctacgggtctcagcgtcga	104832

RESULT	14			
AX147772/c				
LOCUS	AX147772	801 bp	DNA	PAT 08-JUN-2001
DEFINITION	Sequence 17 from Patent WO0136473.			
AX147772				
VERSION	AX147772.1	GI:14346807		
KEYWORDS				
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryote; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
AUTHORS	1 (bases 1 to 801)			
	Vogel, G., Wood, L.S., Parodi, L.A., Hiebsch, R.R., Lind, P., Slightom, J., Schellin, K.A., Kayes, P.S., Bannigan, C.M., Ruff, V., Sejltz, T. and Huff, R.M.			
TITLE	Novel g protein-coupled receptors			
JOURNAL	Patent: WO 0136473-A 17 25-MAY-2001;			
FEATURES	PHARMACIA and UPJOHN COMPANY (US)			
source	Location/Qualifiers			
	1..801			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
BASE COUNT	200 a 187 c 261 g 153 t			
ORIGIN				

Query Match	38.4%	Score 472.6	DB 6	Length 801
Best Local Similarity	79.5%	Pred. No. 4.2e-120		
Matches 578	Conservative 0	Mismatches 134	Indels 15	Gaps 1
QY	17	tcagtaatcgggaaacttgaaatgtcttccttgatccac-----gatcct	61	
DB	801	TCAGGATGGAANAACCTTCAGATCCTTCCTGGATCTCCACGACGANAACCTAGACATCCA	742	
QY	62	ctcaatgaatctatgaaacagcacagaggaatgtaactggtcccaactctgttgcaccaaagcgc	121	
DB	741	TTCCGAAACACACTGACACGACGAGAGATATGCGCCTTCCTGTGGGAGACTGGGGC	682	
QY	122	agtgaactatccctccggtgtctgttgcctatgtcgctgtatcttcctcgtgtggglaatg	181	
DB	681	AGCCACTTCTCTCCGCCGTGTCTGTGGTATATGACCAATTTTGTGTGGGGTCAATT	622	
QY	182	ggcaatctctgtgtgtcatgttgttltgcgaatcgaacttgaagacacccccaac	241	
DB	621	GGCAATGTCCTGGTGTGCTGTGATTCGTGACGACACAGGCTATGAAGACGCCACCAAC	562	
QY	242	tactatctctcagctgtgcaagtccacagatcgcgtgtgtccctctcttgggagtgactctg	301	
DB	561	TACTACTCTTCAGCGCTGGCGGTCTCTGTACCTCTCTGGTCTCTCTTGGAAATGCCCTG	502	
QY	302	gaaatctacgaatgttggcacaaatatacccttctcgttctcgagcctgttggatgtacttc	361	
DB	501	GAGGTCTATGAGATGTGGGGCAATACCTTTCTTGTGGGGCCGTGGGCTGCTAATTC	442	
QY	362	aagacagccctcttcgaactgtgtgtcttcttcacatctcagatgtcaccaaggttgc	421	

Db 441 AACAGCGCCCTCTTTGAGACCGTGTGCTTCGCCCTCCATCCATCGACATCAACACCGCTGACG 382
QY 422 gttagagcgtatgtggtgcatctgtccacccttccgagcgaagtgtagagcgaagcga 481
Db 381 GTGAGACGGCTACGTGGCCATCTTACACCCGTTCCGGCCCAACTGACAGACCCGCGC 322
QY 482 cgggcccctcaagatctccagcagctgtgagcttctctgtgtgtcttcttctcccaat 541
Db 321 CGGGCCCTGAGATCCTCGGCGATGCTGTGGGGCTTCGCCGCTCTCTTCCTCCGCGCCAAC 262
QY 542 accagagatcagatgcatcaagttccagcagacttcccaaggggtccctccgtactgtctca 601
Db 261 ACCAGATCCATCGGATCAAGGCTGCTACCTTCCCAATGGGTCCCTGGTCCAGGTTTCG 202
QY 602 gcaacttgcagatgccaacacacatgtgtgtatgaacttgatccaagctaccagc 661
Db 201 GCCACCTGTACGGTTCATCAAGCCCATGTGATCTACATTTTCATCCAGGTCACCTCC 142
QY 662 ttcctctctacatctcccaatgaccctcatcagcgtccctctactactcatctatgtggctc 721
Db 141 TTCTTATTTCTACCTCCTCCCATGACTGATCATGATGTCTCTTACTACTGATGACATC 82
QY 722 aggtctga 728
Db 81 AGAGTGA 75

RESULT 15

AX139107 AX139107 729 bp DNA PAT 30-MAY-2001
LOCUS AX139107
DEFINITION Sequence 1 from Patent EP1090990.
ACCESSION AX139107
VERSION AX139107.1 GI:14274786
KEYWORDS

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Granulata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE

AUTHORS

TITLE

JOURNAL

Human g-protein-coupled receptor

Patent: EP 1090990-A 1 11-APR-2001;

Pfizer Limited (GB); PFIZER INC. (US)

Location/Qualifiers

1. 729

/organism="Homo sapiens"

/db_xref="taxon:9606"

BASE COUNT 135 a 245 c 167 g 182 t

ORIGIN

Query Match

38.2%; Score 469.8; DB 6; Length 729;

Best Local Similarity

79.6%; Pred. No. 2.5e-119;

Matches 574:

Conservative 0; Mismatches 132; Indels 15; Gaps 1;

Db 241 CTCTTAGCGCTGGCGGCTCTCTGACCTCTGTCCTGCTCTCTCTGGAATGCCCGCTGAGGTC 300
QY 308 taccagatctgagacataataaccttccctggttgcggccgttgagagctacttcaagaca 367
Db 301 TATGAGATGTGGGCACTACCTCTTGTGTGGGCGCGGTGGCTGTACTACTCAAGAGC 360
QY 368 gcccttccgagactgtgtgttctgctccatctcagttcagttccacaggtttagcgtagag 427
Db 361 GCCCTCTTGGAGACCGTGTGCTTCGGCTCCATCTCAGCATGACACCGGTGAGGAG 420
QY 428 cgtatgtgacatgtgtccacccttccgagcgaagcttgagagcagcggcgagagcc 487
Db 421 CGCTACGTGGCCATCCCTACACCGGTTCCGGCCAAATGACAGACCCGGCGCCG 480
QY 488 cttagatctccagcctagctgtgagcttctctgtgtgtcttcttcttgcacaaataccagc 547
Db 481 CTGAGATCTCTGGCATCTGCTGTGGGCTTCTCGGTCTTCTCCCTGGCCAAACACACG 540
QY 548 atccatgagcatcaagttcagcagcatttcccaacgggttctccgtatcgttcaagcacc 607
Db 541 ATCCATGGCATCAAGTTCCACTACTTCCCAATGGGTCCCTGGTCCAGGTTGGCCACC 600
QY 608 tgcacagttcccaaacccatgtgtgtatgaactgtatcatccaagctaccagcttctcc 667
Db 601 TGTACGTCATCAAGCCCATGTGATCTACAAATTTATCATCAAGGTCACCTCTTCCGA 660
QY 668 ttctacatctcccaatgagacctcatcagcgttctctactactcaatgtgggtcagagctg 727
Db 661 TTCTACTCTCTCCCATGACTGATCATGATGTCTCTTACTACTCATGGCATTCAAGAGT 720
QY 728 A 728
Db 721 A 721

Search completed: April 22, 2002, 18:22:55
Job time: 11:34 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2002, 14:13:42 ; Search time 43.26 Seconds
(without alignments)
1403.214 Million cell updates/sec

Title: US-09-609-146-4

Perfect score: 2185
Sequence: 1 MSOMEXLQNSMYQCKLED.....ALSSQMSRRTVQSFHNKT 415

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: SPREMBL_17:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mhc:*
8: sp_mammal:*
9: sp_organelle:*
10: sp_plant:*
11: sp_protist:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2185	100.0	415	4	09GZ04
2	2170	99.3	412	4	09NRA6
3	1650.5	75.5	395	11	09JTB1
4	1643.5	75.2	395	11	09ES04
5	942	43.1	403	4	043664
6	942	43.1	426	4	09HB89
7	885.5	40.5	402	11	09JTB2
8	880.5	40.3	412	11	09JTB5
9	825.5	37.8	405	11	05S040
10	552.5	25.3	428	5	09VEN4
11	515	23.6	595	5	09VFW6
12	497.5	22.8	660	5	09VFW5
13	480	22.0	378	5	018701
14	470.5	21.5	418	5	017239
15	469	21.5	363	13	093413
16	444.5	20.3	374	13	093412
17	424	19.4	397	13	09DDR1
18	390	17.8	401	13	09DDR0
19	361.5	17.5	404	13	09DFB0

20	378	17.3	559	13	093414	093414 spherooides
21	376	17.2	513	13	09DFA9	09df9 catostomus
22	371.5	17.0	377	13	098U14	098u14 brachydanio
23	361	16.5	383	13	042324	042324 catostomus
24	360	16.5	438	11	09R0D1	09r0d1 mus musculu
25	356.5	16.3	352	11	088820	088820 rattus norv
26	356.5	16.3	352	11	09R297	09r297 rattus norv
27	356.5	16.3	352	11	09QW3	09qw3 rattus norv
28	356.5	16.3	452	11	09JKN0	09jkn0 mus musculu
29	355.5	16.3	393	11	09R1M0	09r1m0 mus musculu
30	355.5	16.3	401	11	09R1L9	09r1l9 mus musculu
31	355.5	16.3	444	11	09JTY1	09jty1 mus musculu
32	355	16.2	454	4	09H573	09h573 homo sapien
33	354	16.2	400	6	09MYW9	09myw9 macaca mula
34	351	16.1	464	5	09GQ54	09gq54 aedes aegypt
35	348	15.9	504	5	09VGX8	09vgx8 drosophila
36	347	15.9	394	5	09U721	09u721 drosophila
37	346.5	15.9	367	13	09PVG0	09pvg0 carassius a
38	346.5	15.9	367	13	09PVP9	09pvp9 carassius a
39	346	15.8	384	13	098UH1	098uh1 brachydanio
40	346	15.8	411	13	09M6T3	09m6t3 gallus gall
41	345	15.8	388	4	09U1Y1	09u1y1 homo sapien
42	341.5	15.6	398	11	09JKA0	09jka0 mus musculu
43	339	15.5	398	4	09UDE7	09ude7 homo sapien
44	336	15.4	429	4	013729	013729 homo sapien
45	336	15.4	455	4	060451	060451 homo sapien

ALIGNMENTS

RESULT	ID	PRELIMINARY:	PRT:	415 AA.
09GZ04	09GZ04	09GZ04		
AC	01-MAR-2001 (TREMBLrel. 16, Created)			
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)			
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)			
DE	NEUROMEDIN U RECEPTOR 2 (NEUROMEDIN U RECEPTOR-TYPE 2).			
GN	NMUR2.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
NCBI	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE-20490668; PubMed-10899166;			
RA	Raddatz R., Wilson A.E., Artymyshyn R., Bonini J.A., Borowsky B.,			
RA	Boteju L.W., Zhou S., Kouranova E.V., Ngorny R., Guevarra M.S.,			
RA	Dal M., Letman G.S., Vayse P.J., Branchek T.A., Gerald C., Forray C.,			
RA	Adham N.;			
RT	"Identification and Characterization of Two Neuromedin U Receptors			
RT	Differentially Expressed in Peripheral Tissues and the Central Nervous			
RT	System.";			
RL	J. Biol. Chem. 275:32452-32459(2000).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	PubMed-11010960;			
RA	Shan L., Qiao X., Crona J.H., Behan J., Wang S., Laz T., Bayne M.,			
RA	Gustafson E.L., Monsma F.J. Jr., Hedrick J.A.;			
RT	"Identification of a Novel Neuromedin U Receptor Subtype Expressed in			
RT	the Central Nervous System.";			
RL	J. Biol. Chem. 275:39482-39486(2000).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RA	Pang L., Wang S., Laz T., Hedrick J.A.;			
RL	Submitted (Aug-2000) to the EMBL/Genbank/DBJ databases.			
DR	EMBL, AF272363; AAC24794.1; -			
DR	EMBL, AF292402; AAC03064.1; -			
DR	InterPro: IPR000276; GPCR_Rhodopsn.			
DR	Pfam: PF00001; 7tm_1; 1.			
DR	PRINTS: PR00237; GPCRHOPOPSN.			
DR	PROSITE: PS00237; G_PROTEIN_RECPT_FL_1; UNKNOWN_1.			

DR PROSITE: PS50262; G_PROTEIN_RECEP_F1_2; 1.
 KM Receptor.
 SQ SEQUENCE 415 AA: 47725 MW: C2BACD84B313390F CRC64:

Query Match 100.0%; Score 2185; DB 4; Length 415;
 Best Local Similarity 100.0%; Pred. No. 2.5e-154;
 Matches 415; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MSGMEKLNASWYQOKLEDPEFKHLNSTEYLAFLCGPGRSHFFLPVSVVYPIFVVG 60
 DB 1 MSGMEKLNASWYQOKLEDPEFKHLNSTEYLAFLCGPGRSHFFLPVSVVYPIFVVG 60
 OY 61 IGVNLVCLVILQHQAMKPTNNYLFSLAVSDLVLLGMPLEVEYEMRNYPFLFGVGCY 120
 DB 61 IGVNLVCLVILQHQAMKPTNNYLFSLAVSDLVLLGMPLEVEYEMRNYPFLFGVGCY 120
 OY 121 FKTALFETVCFASILSTTVSVERYVAIIHPRAKQSTRRRRLRLGLVWGSVLSPLP 180
 DB 121 FKTALFETVCFASILSTTVSVERYVAIIHPRAKQSTRRRRLRLGLVWGSVLSPLP 180
 OY 181 NTSIGIKFHPNGSLVGSATCTVIKPMWYINFTIOYTSFLFYLLPMTVISLVYLA 240
 DB 181 NTSIGIKFHPNGSLVGSATCTVIKPMWYINFTIOYTSFLFYLLPMTVISLVYLA 240
 OY 241 LRLKDKSLAEDEGNANIOPCRKSYNKMFLVLVFAICWAPFHIDRLFFSFVEEMSES 300
 DB 241 LRLKDKSLAEDEGNANIOPCRKSYNKMFLVLVFAICWAPFHIDRLFFSFVEEMSES 300
 OY 301 LAAVNLVHVSGVFYLSAVNPITYNLRSRFOAFQNVISSEFKQMSQHPOLPPA 360
 DB 301 LAAVNLVHVSGVFYLSAVNPITYNLRSRFOAFQNVISSEFKQMSQHPOLPPA 360
 OY 361 QNIFTECHFVELTEDIQPOFPCOSSMHNHSLPTALSSQMSRTYQSFHFNKT 415
 DB 361 QNIFTECHFVELTEDIQPOFPCOSSMHNHSLPTALSSQMSRTYQSFHFNKT 415

RESULT 2

O9NRA6 PRELIMINARY: PRT: 412 AA.

AC O9NRA6: 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE NEUROMEDIN U RECEPTOR 2 (G PROTEIN-COUPLED RECEPTOR TGR-1).
 GN NM2R OR TGR-1.
 OS Homo sapiens (Human).
 OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20351041; PubMed=10894543;
 RA Howard A.D., Wang R., Pong S.-S., Mellin T.N., Strack A., Guan X.-M.,
 RA Zeng Z., Williams D.L., Feiguer S.D., Nunes C.N., Murphy B.,
 RA Stal J.N., Yu H., Jiang Q., Clements M.K., Tan C.P., McKee K.K.,
 RA Hrenluk D.L., McDonald T.P., Lynch K.R., Evans J.F., Austin C.P.,
 RA Caskey T., van der Ploeg L.H.T., Liu Q.;
 RT "Identification of receptors for neuromedin U and its role in
 feeding.";
 RL Nature 406:70-74(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX PubMed=10887190;
 RA Hosoya M., Moriya T., Kawamata Y., Ohkubo S., Fujii R., Matsui H.,
 RA Shintani Y., Fukusumi S., Habata Y., Hinuma S., Onda H., Nishimura O.,
 RA Fujino M.;
 RT "Identification and Functional Characterization of a Novel Subtype of
 Neuromedin U Receptor.";
 RL J. Biol. Chem. 275:29528-29532(2000).
 DR EMBL: AF242874; AAF82755.1; -;
 DR EMBL: AB041228; BAB13721.1; -;

DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR Pfam: PF00001; 7tm_1; 1.
 DR PRINTS: PR00237; GPCRHHODPSN.
 DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
 DR PROSITE: PS50262; G_PROTEIN_RECEP_F1_2; 1.
 KM Receptor.
 SQ SEQUENCE 412 AA: 47450 MW: 33306D14A841B63E CRC64:

Query Match 99.3%; Score 2170; DB 4; Length 412;
 Best Local Similarity 100.0%; Pred. No. 3.2e-153;
 Matches 412; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 MEKLNASWYQOKLEDPEFKHLNSTEYLAFLCGPGRSHFFLPVSVVYPIFVVG 63
 DB 1 MEKLNASWYQOKLEDPEFKHLNSTEYLAFLCGPGRSHFFLPVSVVYPIFVVG 60
 OY 64 VLVCLVILQHQAMKPTNNYLFSLAVSDLVLLGMPLEVEYEMRNYPFLFGVGCY 123
 DB 61 VLVCLVILQHQAMKPTNNYLFSLAVSDLVLLGMPLEVEYEMRNYPFLFGVGCY 120
 OY 124 ALFETVCFASILSTTVSVERYVAIIHPRAKQSTRRRRLRLGLVWGSVLSPLP 183
 DB 121 ALFETVCFASILSTTVSVERYVAIIHPRAKQSTRRRRLRLGLVWGSVLSPLP 180
 OY 184 IHGIRKHPNGSLVGSATCTVIKPMWYINFTIOYTSFLFYLLPMTVISLVYLA 243
 DB 181 IHGIRKHPNGSLVGSATCTVIKPMWYINFTIOYTSFLFYLLPMTVISLVYLA 240
 OY 244 KDKSLAEDEGNANIOPCRKSYNKMFLVLVFAICWAPFHIDRLFFSFVEEMSESLA 303
 DB 241 KDKSLAEDEGNANIOPCRKSYNKMFLVLVFAICWAPFHIDRLFFSFVEEMSESLA 300
 OY 304 VFNLVHVSGVFYLSAVNPITYNLRSRFOAFQNVISSEFKQMSQHPOLPPAQRN 363
 DB 301 VFNLVHVSGVFYLSAVNPITYNLRSRFOAFQNVISSEFKQMSQHPOLPPAQRN 360
 OY 364 IFLTECHFVELTEDIQPOFPCOSSMHNHSLPTALSSQMSRTYQSFHFNKT 415
 DB 361 IFLTECHFVELTEDIQPOFPCOSSMHNHSLPTALSSQMSRTYQSFHFNKT 412

RESULT 3

O9JIB1 PRELIMINARY: PRT: 395 AA.

AC O9JIB1: 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE NEUROMEDIN U RECEPTOR 2.
 GN NM2R.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=SPRAGUE-DAWLEY;
 RX MEDLINE=20351041; PubMed=10894543;
 RA Howard A.D., Wang R., Pong S.-S., Mellin T.N., Strack A., Guan X.-M.,
 RA Zeng Z., Williams D.L., Feiguer S.D., Nunes C.N., Murphy B.,
 RA Stal J.N., Yu H., Jiang Q., Clements M.K., Tan C.P., McKee K.K.,
 RA Hrenluk D.L., McDonald T.P., Lynch K.R., Evans J.F., Austin C.P.,
 RA Caskey T., van der Ploeg L.H.T., Liu Q.;
 RT "Identification of receptors for neuromedin U and its role in
 feeding.";
 RL Nature 406:70-74(2000).
 RN EMBL: AF242875; AAF82756.1; -;
 DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR Pfam: PF00001; 7tm_1; 1.
 DR PRINTS: PR00237; GPCRHHODPSN.
 DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
 DR PROSITE: PS50262; G_PROTEIN_RECEP_F1_2; 1.

RW Receptor. 395 AA; 44722 MW; 01D3765B5D5355C0 CRC64;
SQ SEQUENCE

Query Match 75.5%; Score 1650.5; DB 11; Length 395;
Best Local Similarity 79.4%; Pred. No. 9.4e-115;
Matches 312; Conservative 36; Mismatches 40; Indels 5; Gaps 1;

```
OY 4 MEKLNASWIIYQOKLEDPQKHLSNTEEYLAFICGPRSHFFLPVSVVYPIFVGVIGN 63
DB 1 MCKLENASWIIH-----DPLMKYLNSTEEYLAHLGCPKRSDSLPSVAVALITLVGMGN 55
OY 64 VLVCVLILQHOAKMKTPTNYLYFLSLAVSDLLVLLGMPLEVEYEMRNYPFLFGVGCYFKT 123
DB 56 LVCVYIVRHQTKPTNYLYFLSLAVSDLLVLLGMPLEIYEMNHNYPFLFGVGCYFKT 115
OY 124 ALPEYVCFASILSTTVSVRYVAIIHPRAKQSTRRALRLILGIWGSVLSPLPNTS 183
DB 116 ALPEYVCFASILSTTVSVRYVAIIHPRAKLESTRRALRLILSLWSESVLSPLPNTS 175
OY 184 IHGIRFHPNGSLVPGSATCTVYKPMIYNFIQVTSFLFYLLPMTVISLVLYMALRL 243
DB 176 IHGIRFHPNGSSVPGSATCTVYKPMIYNFIQVTSFLFYLLPMTVISLVLYMALRL 235
OY 244 KKDLSLEADGEGNANIORPCRSYVNMKLFVLVFAICMAFPHIDRLFFSFVEESESIAA 303
DB 226 KRDESLKANKVAVNIHPRSKSVTKMLFVLVFAICMFPFHIDRLFFSFVEESESIAA 295
OY 304 VFNLVHVYSGVFYLSAVNPITYNLSRFOAFQNVISFHKQMSHDPQLPPAQRN 363
DB 296 VFNLIHVYSGVFYLSAVNPITYNLSRFRRAAFRNVSPTCKWCHPRHQPGPAOKI 355
OY 364 IFLTECHFEVLTEDDIGPOPCOSSMHNHSLPTA 396
DB 356 IFLTECHFEVLTEDDAGPOFGOSSIHNTNLPTA 388
```

RESULT 4
OQESQ4 PRELIMINARY; PRT; 395 AA.
AC OQESQ4:
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE G PROTEIN-COUPLED RECEPTOR TGR-1.
GN TGR-1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=10887190;
RA Hosoya M., Moriya T., Kawamata Y., Ohkubo S., Fujii R., Matsui H.,
RA Shintani Y., Fukusumi S., Habata Y., Hinuma S., Onda H., Nishimura O.,
RA Fujino M.;
RT "Identification and Functional Characterization of a Novel Subtype of
RT Neurexin-1 Receptor";
RL J. Biol. Chem. 275:29528-29532(2000).
DR EMBL; AB041229; BAB13722.1;
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm.1.1.
DR PRINTS; PR00237; GPCR_Rhodopsn.
DR PROSITE; PS00237; G_PROTEIN_RECPEP_F1_1; UNKNOWN_1.
DR PROSITE; PS50262; G_PROTEIN_RECPEP_F1_2; 1.
RW Receptor.
SQ SEQUENCE 395 AA; 44756 MW; EABD6D36E9C355DA CRC64;

Query Match 75.2%; Score 1643.5; DB 11; Length 395;
Best Local Similarity 78.9%; Pred. No. 3.1e-114;
Matches 310; Conservative 37; Mismatches 41; Indels 5; Gaps 1;

```
OY 4 MEKLNASWIIYQOKLEDPQKHLSNTEEYLAFICGPRSHFFLPVSVVYPIFVGVIGN 63
DB 1 MCKLENASWIIH-----DPLMKYLNSTEEYLAHLGCPKRSDSLPSVAVALITLVGMGN 55
OY 64 VLVCVLILQHOAKMKTPTNYLYFLSLAVSDLLVLLGMPLEVEYEMRNYPFLFGVGCYFKT 123
DB 56 LVCVYIVRHQTKPTNYLYFLSLAVSDLLVLLGMPLEIYEMNHNYPFLFGVGCYFKT 115
OY 124 ALPEYVCFASILSTTVSVRYVAIIHPRAKQSTRRALRLILGIWGSVLSPLPNTS 183
DB 116 ALPEYVCFASILSTTVSVRYVAIIHPRAKLESTRRALRLILSLWSESVLSPLPNTS 175
OY 184 IHGIRFHPNGSLVPGSATCTVYKPMIYNFIQVTSFLFYLLPMTVISLVLYMALRL 243
DB 176 IHGIRFHPNGSSVPGSATCTVYKPMIYNFIQVTSFLFYLLPMTVISLVLYMALRL 235
OY 244 KKDLSLEADGEGNANIORPCRSYVNMKLFVLVFAICMAFPHIDRLFFSFVEESESIAA 303
DB 226 KRDESLKANKVAVNIHPRSKSVTKMLFVLVFAICMFPFHIDRLFFSFVEESESIAA 295
OY 304 VFNLVHVYSGVFYLSAVNPITYNLSRFOAFQNVISFHKQMSHDPQLPPAQRN 363
DB 296 VFNLIHVYSGVFYLSAVNPITYNLSRFRRAAFRNVSPTCKWCHPRHQPGPAOKI 355
OY 364 IFLTECHFEVLTEDDIGPOPCOSSMHNHSLPTA 396
DB 356 IFLTECHFEVLTEDDAGPOFGOSSIHNTNLPTA 388
```

RESULT 5
OQ43664 PRELIMINARY; PRT; 403 AA.
AC OQ43664:
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE ORPHAN G PROTEIN-COUPLED RECEPTOR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tan C.T., McKee K.K., Palyha O.C., Feighner S.D., Hreniuk D.L.,
RA Smith R.G., Howard A.D.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF044601; AAC02680.1;
DR EMBL; AF044600; AAC02680.1; JOINED.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm.1.1.
DR PRINTS; PR00237; GPCR_Rhodopsn.
DR PROSITE; PS00237; G_PROTEIN_RECPEP_F1_1; UNKNOWN_1.
DR PROSITE; PS50262; G_PROTEIN_RECPEP_F1_2; 1.
SQ SEQUENCE 403 AA; 44980 MW; CF8D5F1A961CCBD9 CRC64;

Query Match 43.1%; Score 942; DB 4; Length 403;
Best Local Similarity 50.8%; Pred. No. 2.9e-62;
Matches 180; Conservative 63; Mismatches 89; Indels 22; Gaps 6;

```
OY 20 DPOQKHLSNTEEYLAFICGPRSHFFLPVSVVYPIFVGVIGNVLCVILQHOAKMKT 78
DB 13 DP--EDNLNLTDEALRLKYLGAPOQTELPMPICATYLLFVVGAVNGTCLVILRHAKMKT 70
OY 79 PTNYLYFLSLAVSDLLVLLGMPLEVEYEMRNYPFLFGVGCYFKTALFETVCFASILST 138
DB 71 PTNYLYFLSLAVSDLLVLLGMPLEIYEMNHNYPFLFGVGCYFKTLLFEVAVLASVLTNT 130
OY 139 TVSVRYVAIIHPRAKQSTRRALRLILGIWGSVLSPLPNTSIGHGIRFHPNGSLV 198
DB 131 ALSVERYAVAVHQLQARSMVTRAHVRYLGAVGLAMLSLPTSLHGIRFHPNGRVP 190
OY 199 PGSATCTVYKPMIYNFIQVTSFLFYLLPMTVISLVLYMALRLKDKLSLEADGEGN--- 255
```

Db 191 PSASVCMVLRPRALYNNVQTALLFFCLPMALMSVLYLLIGRLRRERILLMQEAKGRG 250
Oy 256 ----ANTQRC-----RKSVMKMLFVLVYVAICWAPFHIDRLEFFSYEEMSESIAV 304
Db 251 SAASRSRYTCRLQOHDGRGRQVTKMLFVLVYVVGICWAPFHADRVMSVVSQMTDGLHLA 310
Oy 305 FNLVHVVSQVFYLSAVNPDIYNILSRROAFQWVI---SSFH--KQWHSQH 353
Db 311 FQHVHVISGIFFLGSAANPVLYLSMSRFRFQDALCLGACCHRLRPHSSH 364

RESULT 6
O9HB89 PRELIMINARY: PRT: 426 AA.

AC O9HB89:
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE NEUROMEDIN U RECEPTOR 1.
GN NMUR1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20490668; PubMed=10899166;
RA Raddatz R., Wilson A.E., Artyushyn R., Bonini J.A., Borowsky B.,
Bojcu L.W., Zhou S., Kouranova E.V., Nagorny R., Guevarra M.S.,
Rah D.M., Lerman G.S., Vayse P.J., Branchek T.A., Gerald C., Fortay C.,
Adham N.,
RT "Identification and Characterization of Two Neuremodin U Receptors
RT Differentially Expressed in Peripheral Tissues and the Central Nervous
RT System."
RL J. Biol. Chem. 275:32452-32459(2000).
DR EMBL: AF272362; AAG24793.1; -.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PR00237; GPCR_RHODOPSN.
DR PROSITE: PS00237; G_PROTEIN_RECEPTOR_FL1; UNKNOWN_1.
DR PROSITE: PS50262; G_PROTEIN_RECEPTOR_FL2; 1.
KW Receptor.
SQ SEQUENCE 426 AA: 47350 MW: F8DD781C020F04AA CRC64;

Query Match 43.1%; Score 942; DB 4; Length 426;
Best Local Similarity 50.8%; Pred. No. 3.1e-62;
Matches 180; Conservative 63; Mismatches 89; Indels 22; Gaps 6;

Oy 20 DFOKHLNSTEYLAF-LCGPRRSHFPLPVSVYVPIFVGVIGNVLCVLIHQHQAAMKT 78
Db 36 DF--EDLNTLDEALRLKYLIGPQOTELFMPICATYLLIFVGVANGNLCTCYILRHKAMKT 93
Oy 79 PNNYLFSLASDGLVLLGLMPLVEYEMKRNPFLEPGVGCYFKYLLFETVCASILST 138
Db 94 PNNYLFSLASDGLVLLGLMPLVEYEMKRNPFLEPGVGCYFKYLLFETVCASILST 133
Oy 139 TVSEYEVVALHPERAKLOSTRRALRLIGIVGFSVLSPTSHIGIKFHFNPQSLV 198
Db 154 ALSVEYEVVALHPERAKLOSTRRALRLIGIVGFSVLSPTSHIGIKFHFNPQSLV 193
Oy 199 PGSACTVTKPMWYNFIQVTSFLYLLPMTVISVLYLMLALRLKDKSLDEAGN--- 255
Db 214 PSASVCMVLRPRALYNNVQTALLFFCLPMALMSVLYLLIGRLRRERILLMQEAKGRG 273
Oy 256 ----ANTQRC-----RKSVMKMLFVLVYVAICWAPFHIDRLEFFSYEEMSESIAV 304
Db 274 SAASRSRYTCRLQOHDGRGRQVTKMLFVLVYVVGICWAPFHADRVMSVVSQMTDGLHLA 333
Oy 305 FNLVHVVSQVFYLSAVNPDIYNILSRROAFQWVI---SSFH--KQWHSQH 353
Db 334 FQHVHVISGIFFLGSAANPVLYLSMSRFRFQDALCLGACCHRLRPHSSH 387

RESULT 7
ID O9JIB2 PRELIMINARY: PRT: 402 AA.

AC O9JIB2:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE NEUROMEDIN U RECEPTOR 1.
GN NMUR1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY;
RX MEDLINE=20351041; PubMed=10894543;
RA Howard A.D., Wang R., Pong S.-S., Mellin T.N., Strack A., Guan X.-M.,
Zeng Z., Williams D.L., Feighner S.D., Nunes C.N., Murphy B.,
Stall J.N., Yu H., Jiang Q., Clements M.K., Tan C.P., McKee K.K.,
Hreniuk D.L., McDonald T.P., Lynch K.R., Evans J.F., Austin C.P.,
RA Caskey T., van der Ploeg L.H.T., Liu Q.,
RT "Identification of receptors for neuremodin U and its role in
RT feeding."
RL Nature 406:70-74(2000).
DR EMBL: AF242873; AAF82754.1; -.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PR00237; GPCR_RHODOPSN.
DR PROSITE: PS00237; G_PROTEIN_RECEPTOR_FL1; UNKNOWN_1.
DR PROSITE: PS50262; G_PROTEIN_RECEPTOR_FL2; 1.
KW Receptor.
SQ SEQUENCE 402 AA: 45782 MW: CB576083D9F7A095 CRC64;

Query Match 40.5%; Score 885.5; DB 11; Length 402;
Best Local Similarity 48.3%; Pred. No. 4.5e-58;
Matches 173; Conservative 56; Mismatches 108; Indels 21; Gaps 4;

Oy 19 EDPQKNI-----LNSTEYLAF-LCGPRRSHFPLPVSVYVPIFVGVIGNVLCVLIHQ 72
Db 5 DSEKHEFDLEDNLTHEDRLKYLIGPQVQKQFPLCYLLIFVGVTLGNGLCTCYILR 64
Oy 73 HOAKMPTNYLFSLASDGLVLLGLMPLVEYEMKRNPFLEPGVGCYFKYLLFETVCFA 132
Db 65 QKMHFTNYLFSLASDGLVLLGLMPLVEYEMKRNPFLEPGVGCYFKYLLFETVCFA 124
Oy 133 SILSTTVSEYEVVALHPERAKLOSTRRALRLIGIVGFSVLSPTSHIGIKFHF 192
Db 125 SVLNTVALSEYEVVALHPERAKLOSTRRALRLIGIVGFSVLSPTSHIGIKFHF 184
Oy 193 PNSLVPGSATCTYIKRMWYNFIQVTSFLYLLPMTVISVLYLMLALRLKDKSLDEAD 252
Db 185 PCGPPVDVSTCYLRQPFYKLVITQTLFFCLPMWTSVLYLLIGRLRRERILLQ 244
Oy 253 EGNANTQRC-----RKSVMKMLFVLVYVAICWAPFHIDRLEFFSYEEMSES 298
Db 245 EVKGRISAARQASHRSIQLDRERROYTKMLALYVFCGCVWPFHADRLMSMVSHT 304
Oy 299 ESLAAVFNLVHVVSQVFYLSAVNPDIYNILSRROAFQWVI---SSFH--KQWHSQH 356
Db 305 DGLRLAQSVHLASGVFLYLSAANPVLYLSMSRFRFQDALCLGACCHRLRPHSSH 361

RESULT 8
ID O9JIB2 PRELIMINARY: PRT: 412 AA.

AC O9JIB2:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegyan C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Matrei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Moberly C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Piltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of *Drosophila melanogaster*."
 RL Science 287:2185-2195(2000).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 DR EMBL: AE003703; AAF55016.1; -
 DR Flybase: FBgn0038201; CG9918.
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR Pfam: PF00001; 7tm_1; 1.
 DR PRINTS: PR00237; GPCRHOOPS.
 DR PROSITE: PS00237; G-PROTEIN_RECEP_FL1; 1.
 DR PROSITE: PS50262; G-PROTEIN_RECEP_FL2; 1.
 KW G-protein coupled receptor; Glycoprotein; Transmembrane.
 SQ SEQUENCE 428 AA: 47198 MW: 9DDA41E93046BBD7 CRC64;

Query Match 25.3%; Score 552.5; DB 5; Length 428;
 Best local similarity 33.2%; Pred. No. 2, 2e-33;
 Matches 121; Conservative 73; Mismatches 98; Indels 73; Gaps 7;

38 GPRRS--HFLPVSVYVPPIFVGVGNVLVCLVILHQAAMKPTNYVFLSLAVSDLLV 95
 11 GPRDPLAIVIPYVYVSLIFIGVGNISTCIVIKKRNMTATNYVFLSLAISDFLL 70
 96 LIGMPLEVEYEMNRNYPLFPVGCYFETALFENYCFASISITTVSEYRVAILHPRAK 155
 71 LSGVPOEVSTWSKYRPFVEEYICIGGLAETSANATVLTITAFYERITAIHPPLGQ 130
 156 LQSTRRAALILIGVGSVLSLPNTSINGIFHFYFPNGSLVPGSATCTVIRPMWLYNF 215
 131 AMSKLSRAIRIIVLWIMATVTAIPQAQIGIE-HY-----SGVQGCIVAVYVHSPF 182
 216 IIVTSTSLFLPLPVYISVLYIMALRLKDKSLDEADGANIQRPCR----- 263
 183 --QLSTFFELAPMSITLVLLIGVILRSTLV---EGPASVARRQLKSVSPDITLYR 237
 264 -----KSVNKLMLVLVLAICW 281
 238 YGSGSTAMSGNGGSGAGTGLMGSGAQLSVRGRLNHYGTRVLMALVAVVCFLCW 297
 282 APFHIDRLFPSEVSESESL-----AAVFNLVHVSQVFFVLSAVNPIYNLSSRFQAA 337
 298 APFHAQRLIIVAPARGAKLRDQHEFYVTVYVSGVLYLSTCINPLNLNINSHKREAA 357
 338 FQAVT 342
 358 FKAVL 362

DB 358 FKAVL 362

RESULT 11
 O9VFW6 PRELIMINARY; PRT; 595 AA.

AC O9VFW6;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE CG8795 PROTEIN.
 GN CG8795.
 OS *Drosophila melanogaster* (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephygroidae; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Ceoliker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazel R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Aghayani A., An H.-J., Andrews-Frannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegyan C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Matrei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Moberly C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Piltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of *Drosophila melanogaster*."
 RL Science 287:2185-2195(2000).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 DR EMBL: AE003699; AAF54929.1; -
 DR Flybase: FBgn0038139; CG8795.
 DR InterPro: IPR000923; GPCR_blue1.
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR Pfam: PF00001; 7tm_1; 1.
 DR PRINTS: PR00237; GPCRHOOPS.
 DR PROSITE: PS00196; COOPER_BLUE; UNKNOWN_1.
 DR PROSITE: PS00237; G-PROTEIN_RECEP_FL1; 1.
 DR PROSITE: PS50262; G-PROTEIN_RECEP_FL2; 1.
 KW G-protein coupled receptor; Glycoprotein; Transmembrane.
 SQ SEQUENCE 595 AA: 65091 MW: 3EBE183BADF5471E CRC64;

Query Match 23.6%; Score 515; DB 5; Length 595;
 Best local similarity 35.6%; Pred. No. 1, 8e-30;
 Matches 110; Conservative 66; Mismatches 95; Indels 38; Gaps 9;

OY 48 VSVYVPIFVGVGNVLVCLVILHQAAMKPTNYVFLSLAVSDLLVLLIGMPLEVEYEM 107


```

DB 65 LSVGALFIAGVGLNLTICIVISRNNEHMTATNFYLFALISDMILLCGMPQDLYNLM 124
OY 108 R--NRPFLGPGCYCFKRALFETVCFASLTSTTVSVENVYAILHPFRAKLOSTRRAIR 165
DB 125 HPDNP--FSDSICILESVLSTAAANAVLTITTAFTVEYIAICHPFROHTSKLSRAVK 182
OY 166 ILGIWGFVSLSLNTSNGHGFYFPGSLVPSATFVTKPMYIFITQVTSFLPY 225
DB 183 FFAIWMALALLAP-----QATQFSVNMGM-----GTSTGMNDFFAH--VAVSGFLFF 232
OY 226 LIPMTVISLYLMAIRLKKDKSLADEGNANIQRPC-----RKSVMKLFVLVLV 276
DB 233 GGPMTAICVLYLVIGVKIKRSRLQA-----LPRRCYDYNRGISAQTRVIMLVAVAVA 286
OY 277 FAICAPAFHIDLFSSF-----VEEMSESLAAVNLVHVSGVFYLSAANPIITNLL 330
DB 287 FETICAPAFHQAQVAVGSGTSGIESQWEND--VSILDTYTGVLVFLSTCINPLLYNIM 343
OY 331 SRRFOAFC 339
DB 344 SHKFEAFK 352

RESULT 12
OYFWS PRELIMINARY: PRT: 660 AA.
AC OYFWS5.
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE CG8784 PROTEIN.
GN CG8784.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.C., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H.C., Blazer J.R., Champagne M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abell J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Butova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durkin K.J., Evangeliste C.C., Ferraz C., Ferriere S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Gloeck A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Howlin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
RA Jajali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Maltel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merklov G., Mlshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pauleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klimos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spredling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

```

```

RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Zhu H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL: AE003699; AAF54930.1.-.
DR FlyBase: FBgn0038140; CG8784.
DR InterPro: IPR000923; Copper_blue1.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm.1; 1.
DR PRINTS: PR00237; GPCR_RHODOPSIN.
DR PROSITE: PS00196; COPPER_BLUE; UNKNOWN_1.
DR PROSITE: PS00237; G_PROTEIN_RECIP_FL_1; 1.
DR PROSITE: PS50262; G_PROTEIN_RECIP_FL_2; 1.
DR G-protein coupled receptor; Glycoprotein; Transmembrane.
SQ SEQUENCE 660 AA; 72277 MW; C52D1E7A4CD439 CRC64;

```

Query Match 22.8%; Score 497.5; DB 5; Length 660.
 Best Local Similarity 29.6%; Pred. No. 4e-29;
 Matches 133; Conservative 77; Mismatches 159; Indels 81; Gaps 14;

```

OY 20 DPEOKH----LNSTDEYLAFLCG-----PARRHFL--PVSVYVYVIFVWG 59
DB 62 DKFLTHVAHLNITTEENSLNLTSGNGTNASTMAADSPVDESILTRTALTVCAVAILFVAG 121
OY 60 VIGNVLVCLVLIQHOAKMPTNYVYLSLAVSDLVLLGMPLEVEEMHNP--FLGCV 117
DB 122 VGLNLTICIVISRNNFMHMTATNFYLFENLAVSDLLVSSITPELNLN--YDWMPTFDA 179
OY 118 GGYEFTALEFVCFASLTSTTVSVENVYAILHPFRAKLOSTRRAIRLIGVGSFYLE 177
DB 180 MCIMSGVSEMAANATVLTITTAFTVEYIAICHPFROHTSKLSRAIKFPAIMLAFL 239
OY 178 SLPNSTSIHGIKHPYNGSLVPGSATCYIKPMYINFTIQTSTFLIPPTVSVLY 237
DB 240 ALPQAMQFSVVYQ-----NEGYSCTMEND--FYAHFAVSGFLFEGPMTAICVLV 289
OY 238 LMAIRLKKDK--SLEADGNANIQRPCRSYVNMKLFVLVLFALCMAPFHIDRLFSPF- 293
DB 290 LIGVKIKRSRLQSLPRFIDANRGLMAGRIYIMLVAVAVAFPLCMAPFHQRLMAYG 349
OY 294 ----VEEMSESLAAVNLVHVSGVFYLSAANPIITNLLSRRFOAFCQNT----- 342
DB 350 LNLINIGISRDAFNDFRILDTYTGVLVFLSTCINPLLYNIMSHKFEAFKITLRFQGL 409
OY 343 --SFFHKQ-WHSQHD-----POLPPA-----QRNIFLTCHEVELTE 376
DB 410 ANHHHQSQQHQNHSALLRONGSMRLDPASCYNNNALPEYGYRVVQFRC-----R 463
OY 377 DIGQPFCCSMHNSHLPALSSQWSTRN 406
DB 464 DANHQLSDSDIRTTTTTTTTTINSNSMAAGN 493

RESULT 13
OYFWS PRELIMINARY: PRT: 378 AA.
AC O18701.
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE SIMILAR TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS. NCBI GI: 1055105.
GN C48C5.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilita; Rhabdilitidae;
OC Rhabdilitidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]

```

RESULT	ID		
017239			
ID	017239	PRELIMINARY;	PRT; 418 AA.
AC	017239;		
DT	01-JAN-1998 (TREMBLrel. 05, Created)		
DT	01-JAN-1998 (TREMBLrel. 05, Last sequence update)		
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)		
DE	K10B4.4 PROTEIN.		
GN	K10B4.4.		
OS	Caenorhabditis elegans.		
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;		
OC	Rhabditidae; Pelodermineae; Caenorhabditis.		
OX	NCBI_Taxid=6239;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=BRISTOL N2;		

RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Sammons L., Wohlmann P., Antoniou B.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF025463; AAB71009.1; -;
DR InterPro: IPR000276; GPCR_Rhodopsin.
DR Pfam: PF00001; 7tm_1; 2.
DR PRINTS: PR00237; GPCRHHODPSN.
DR PROSITE: PS00237; G_PROTEIN_RECPT_FL_1; UNKNOWN_1.
DR PROSITE: PSS0262; G_PROTEIN_RECPT_FL_2; 1
DR SEQUENCE 418 AA; 47735 MW; CE0416539CA3BB27 CRC64;

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2002, 14:12:47 ; Search time 21.56 Seconds
(without alignments)
433.157 Million cell updates/sec

Title: US-09-609-146-4
Perfect score: 2185
Sequence: 1 MSGMEKLNASWIVQOKLED.....ALSSSEMSRTMYGSHFNKKT 415

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/pdata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/pdata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/pdata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/pdata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/pdata/2/1aa/6C.COMB.pep:*
6: /cgn2_6/pdata/2/1aa/Backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	448	20.5	353	4	US-09-077-675A-3 Sequence 3, Appl1
2	444	20.3	361	4	US-09-077-675A-8 Sequence 8, Appl1
3	444	20.3	366	4	US-09-077-675A-13 Sequence 13, Appl1
4	436.5	20.0	364	4	US-09-077-675A-16 Sequence 16, Appl1
5	417.5	19.1	353	1	US-08-118-270-45 Sequence 45, Appl1
6	417.5	19.1	353	5	PCT-US93-08528-45 Sequence 45, Appl1
7	410	18.8	416	3	US-08-858-876A-4 Sequence 4, Appl1
8	410	18.8	416	4	US-09-472-880-4 Sequence 4, Appl1
9	405.5	18.6	398	2	US-08-288-663A-1 Sequence 1, Appl1
10	402	18.4	302	4	US-09-077-675A-2 Sequence 2, Appl1
11	398	18.2	302	4	US-09-077-675A-7 Sequence 7, Appl1
12	398	18.2	410	3	US-08-858-876A-2 Sequence 2, Appl1
13	398	18.2	410	4	US-09-472-880-2 Sequence 2, Appl1
14	397	18.2	393	1	US-07-629-1041-3 Sequence 3, Appl1
15	377.5	17.3	398	2	US-08-288-663A-15 Sequence 15, Appl1
16	365	16.7	369	1	US-07-816-283-6 Sequence 6, Appl1
17	365	16.7	369	1	US-08-417-103-6 Sequence 6, Appl1
18	365	16.7	369	1	US-08-417-103-16 Sequence 16, Appl1
19	358	16.4	369	2	US-08-411-859-3 Sequence 9, Appl1
20	358	16.4	369	4	US-08-387-707-9 Sequence 9, Appl1
21	357	16.3	400	4	US-08-188-275A-2 Sequence 2, Appl1
22	356.5	16.3	319	4	US-08-832-399-2 Sequence 2, Appl1
23	356.5	16.3	319	4	US-09-372-498-2 Sequence 2, Appl1
24	356	16.3	369	4	US-08-120-601B-9 Sequence 9, Appl1
25	355	16.2	400	3	US-08-889-108-8 Sequence 8, Appl1
26	355	16.2	400	5	PCT-US94-10358-8 Sequence 8, Appl1
27	352.5	16.1	380	3	US-08-676-351-5 Sequence 5, Appl1

28	352.5	16.1	380	4	US-08-188-275A-5 Sequence 5, Appl1
29	351	16.1	384	3	US-09-071-434-3 Sequence 3, Appl1
30	351	16.1	389	3	US-08-430-286A-7 Sequence 7, Appl1
31	350.5	16.0	376	4	US-08-387-707-17 Sequence 17, Appl1
32	350.5	16.0	380	4	US-08-765-743-2 Sequence 2, Appl1
33	349	16.0	369	1	US-07-816-283-8 Sequence 8, Appl1
34	349	16.0	369	1	US-08-417-103-8 Sequence 8, Appl1
35	348.5	15.9	380	1	US-08-149-093A-7 Sequence 7, Appl1
36	348.5	15.9	380	2	US-08-911-245-7 Sequence 2, Appl1
37	348.5	15.9	380	3	US-08-147-592A-2 Sequence 7, Appl1
38	348.5	15.9	380	4	US-09-510-473-7 Sequence 7, Appl1
39	347.5	15.9	391	1	US-07-816-283-2 Sequence 2, Appl1
40	347.5	15.9	391	1	US-07-816-283-4 Sequence 4, Appl1
41	347.5	15.9	391	1	US-08-417-103-2 Sequence 2, Appl1
42	347.5	15.9	391	1	US-08-417-103-4 Sequence 4, Appl1
43	347.5	15.9	391	1	US-08-417-103-14 Sequence 14, Appl1
44	345.5	15.8	367	3	US-08-676-351-2 Sequence 2, Appl1
45	345.5	15.8	391	4	US-08-120-601B-8 Sequence 8, Appl1

ALIGNMENTS

RESULT 1
US-09-077-675A-3
Sequence 3, Application US/09077675A
Patent No. 6242199
GENERAL INFORMATION:
APPLICANT: Pai, Lee-Yuh
APPLICANT: Feighner, Scott C.
APPLICANT: Howard, Andrew D.
APPLICANT: Pong, Sheng-Shung
APPLICANT: Van Der Ploeg, Leonardus H.T.
TITLE OF INVENTION: RECEPTOR ASSAY
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000, 126 E. Lincoln Ave.
CITY: Rahway
STATE: NJ
COUNTRY: USA
ZIP: 07065-0900
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/077, 675A
FILING DATE: 3-JUN-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Cocuzzo, Anna L.
REGISTRATION NUMBER: 42,452
REFERENCE/DOCKET NUMBER: 19590P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 732-594-1273
TELEFAX: 732-594-4720
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 353 amino acids
TYPE: amino acid
STRANDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-077-675A-3
Query Match 20.5%; Score 448; DB 4; Length 353;

Best Local Similarity 34.3%; Pred. No. 1.8e-34;
Matches 104; Conservative 68; Mismatches 103; Indels 28; Gaps 7;

```
OY 48 VSVVYVPEFVGVGNVLCVLIITQHQAMKTPYNYLFSLAVSDLVLLGMPLEYEMW 107
   | : ||||| : || : : : || : || : || : || : || : || : || : || : ||
Db 33 VTATCAVLFVVGIGNLTMLVSRFRMRTTNLYLSMAFSDLLIFLC-MPDLVRLW 91
OY 108 RNVPFLGPGVCYKFKTALFEVCFASLITSTVTVVERVVALHPFRKLOSTRRALRIL 167
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 92 QYRPMNIGNLCKLFQVSESCYATVLTALSVRFALCPFLRAKVVTKRGVRLVI 151
OY 168 GIWGFVLSLPNTSIHGKIFHYFPNGS-----LVGSACTVYIKPMIYNF 215
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 152 LVIMAVAFCSAGPIFVLGVGE---HDNGTDPDPTNECRATEFAVRSGLLTVM--VW--- 203
OY 216 IIOVTSFLPMTVSVLYIMALKDKSLADEGNANIORPCRSYKMKLFLVYL 275
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 204 -----SSVFFLPVFCCLTVLSLGRKLMRRRGEAAVG--SLRDQNHKQTVKMLAVVF 257
OY 276 VFAICMAPFHIDRLFFS-FVEEMSESLAAVENLVHVVGVGFYLSAVNPITYMLSRF 334
   | : || ||| : || : || : || : || : || : || : || : || : || : || : ||
Db 258 AFICMLPFHYGRYLFKSPSGVETAIQISQYCNLVSVLFYLSAAILNPLINIMSKY 317
OY 335 QAA 337
   : |
Db 318 RVA 320
```

RESULT 2
US-09-077-675A-8
Sequence 8, Application US/09077675A
Patent No. 6242199

GENERAL INFORMATION:

APPLICANT: Pal, Lee-Yuh
APPLICANT: Feighner, Scott C.
APPLICANT: Howard, Andrew D.
APPLICANT: Pong, Sheng-Shung
APPLICANT: Van Der Ploeg, Leonardus H.T.
TITLE OF INVENTION: RECEPTOR ASSAY
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000, 126 E. Lincoln Ave.
CITY: Rahway
STATE: NJ
COUNTRY: USA
ZIP: 07065-0900
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/077,675A
FILING DATE: 3-JUN-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Cocuzzo, Anna L.
REGISTRATION NUMBER: 42,452
REFERENCE/DOCKET NUMBER: 19590P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 732-594-1273
TELEFAX: 732-594-4720
TELEX:
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 361 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: protein
US-09-077-675A-8

Query Match 20.3%; Score 444; DB 4; Length 361;
Best Local Similarity 34.0%; Pred. No. 4.4e-34;
Matches 103; Conservative 68; Mismatches 104; Indels 28; Gaps 7;

```
OY 48 VSVVYVPEFVGVGNVLCVLIITQHQAMKTPYNYLFSLAVSDLVLLGMPLEYEMW 107
   | : ||||| : || : : : || : || : || : || : || : || : || : || : ||
Db 41 VTATCAVLFVVGIGNLTMLVSRFRMRTTNLYLSMAFSDLLIFLC-MPDLVRLW 99
OY 108 RNVPFLGPGVCYKFKTALFEVCFASLITSTVTVVERVVALHPFRKLOSTRRALRIL 167
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 100 QYRPMNIGNLCKLFQVSESCYATVLTALSVRFALCPFLRAKVVTKRGVRLVI 159
OY 168 GIWGFVLSLPNTSIHGKIFHYFPNGS-----LVGSACTVYIKPMIYNF 215
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 160 FVIMAVAFCSAGPIFVLGVGE---NOTDPMDTNECRPTFEAVRSGLLTVM--VW--- 211
OY 216 IIOVTSFLPMTVSVLYIMALKDKSLADEGNANIORPCRSYKMKLFLVYL 275
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 212 -----SSVFFLPVFCCLTVLSLGRKLMRRRGEAAVG--ASLRDQNHKQTVKMLAVVF 265
OY 276 VFAICMAPFHIDRLFFS-FVEEMSESLAAVENLVHVVGVGFYLSAVNPITYMLSRF 334
   | : || ||| : || : || : || : || : || : || : || : || : || : || : ||
Db 266 AFICMLPFHYGRYLFKSPSGVETAIQISQYCNLVSVLFYLSAAILNPLINIMSKY 325
OY 335 QAA 337
   : |
Db 326 RVA 328
```

RESULT 3
US-09-077-675A-13
Sequence 13, Application US/09077675A
Patent No. 6242199

GENERAL INFORMATION:

APPLICANT: Pal, Lee-Yuh
APPLICANT: Feighner, Scott C.
APPLICANT: Howard, Andrew D.
APPLICANT: Pong, Sheng-Shung
APPLICANT: Van Der Ploeg, Leonardus H.T.
TITLE OF INVENTION: RECEPTOR ASSAY
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000, 126 E. Lincoln Ave.
CITY: Rahway
STATE: NJ
COUNTRY: USA
ZIP: 07065-0900
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/077,675A
FILING DATE: 3-JUN-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Cocuzzo, Anna L.
REGISTRATION NUMBER: 42,452
REFERENCE/DOCKET NUMBER: 19590P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 732-594-1273
TELEFAX: 732-594-4720
TELEX:
INFORMATION FOR SEQ ID NO: 13:

RESULT 6
 PCT-US93-08528-45
 : Sequence 45, Application PC/TUS9308528
 : GENERAL INFORMATION:
 : APPLICANT: New York University
 : TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
 : TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
 : NUMBER OF SEQUENCES: 348
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: BROWDY AND NEIMARK
 : STREET: 419 Seventh Street, N.W., Suite 300
 : City: Washington
 : STATE: D.C.
 : COUNTRY: USA
 : Zip: 20004
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: PatentIn Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: PCT/US93/08528
 : FILING DATE: 09-SEP-1993
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 07/943,236
 : FILING DATE: 10-SEP-1992
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Townsend, Kevin G.

RESULT 7
 US-08-858-876A-4
 : Sequence 4, Application US/08858876A
 : Patent No. 6022856
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: Daniel CAPUT
 :
 : APPLICANT: Pascale CHALON
 :
 : APPLICANT: Pascual FERRARA
 :
 : APPLICANT: Vita NATALIO
 :
 : TITLE OF INVENTION: Type 2 Neurotensin Receptor
 :
 : TITLE OF INVENTION: (b)(7)-(R2)
 :
 : NUMBER OF SEQUENCES: 12
 :
 : CORRESPONDENCE ADDRESSES:
 :
 : ADDRESSEE: Jacobson, Price, Holman & Stern, PLLC
 :
 : STREET: 400 Seventh Street
 :
 : City: Washington
 :
 : STATE: D.C.
 :
 : COUNTRY: USA
 :
 : Zip: 20004
 :
 : COMPUTER READABLE FORM:
 :
 : MEDIUM TYPE: Floppy disk
 :
 : COMPUTER: IBM PC compatible
 :
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 :
 : SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
 :
 : CURRENT APPLICATION DATA:
 :
 : APPLICATION NUMBER: US/08/858,876A
 :
 : FILING DATE: 19-SEP-1997
 :
 : CLASSIFICATION: 536

FILING DATE: 10-AUG-1993
APPLICATION NUMBER: 286986/1993
FILING DATE: 16-NOV-1993
APPLICATION NUMBER: 325215/1993
FILING DATE: 22-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: Resnick, David S
REGISTRATION NUMBER: 34,235
REFERENCE/DOCKET NUMBER: 44612
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400.
TELEFAX: 617-523-6440
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 398 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-288-663A-1

Query Match 18.6%; Score 405.5; DB 2; Length 398;
Best Local Similarity 29.9%; Pred. No. 2.1e-30;
Matches 119; Conservative 72; Mismatches 130; Indels 77; Gaps 14;

QY 48 VSVVYVPIPV-VGVIGNVLVCLVLIHQAMKTPNTNYLFSLSVSDLLVL-LGMPV---E 102
DB 27 VVLLVLIICGLGIVGINVNLVMTKHMRTPTNCLVSLAVADMLVLAAGLPNITDS 86
QY 103 VEMMNNYFELFCGVCYCFALFETVCFASILITTVSVSRVALIHPRAKLOSTRR 162
DB 87 IYGSW-----VGYVCCLCITVLOYIGINASSCSITAFIERIALCHPKAOFCTFSR 141
QY 163 ALRILGIWGFVSFLSPNTSHGKIFHPNGLVPGSATC-----TVIKPMWITVNT 216
DB 142 AKKILFVVAFTSLYCMLMFELDLNISTYKDAIVI-----SCGYKISRNYSPYILMDFG 197
QY 217 IQVTSFLFLPMTVTSVLYYLALRL-----KKKKSLEADGNNINOR 260
DB 198 V-----FVVPMLATVTLGFIARILFLNPISDPKRENSKWKNDSTHONTNLNVTNS 231
QY 261 PC-----RKSVMKMLFVLVFAICWAFH---IDRLFFS---FVEEMSESLAAVFNLY 308
DB 252 RCNSTSVSSRKQYTKMLAVVILFALLMMPYRLLVYVNSFLSPQENM-----F 301
QY 309 HVSAGYFFLLSSAVNPITYNLSRRQAQAFQNVISFFHKOMHSQHPOLPRAORNIFLT- 367
DB 302 LFCRICIYLNSAIVNPVIYNLMSQKFRFAFRKLCNCOK-----PTKRPANYSVALNY 354
QY 368 -----ECHFVELTEDIGQFPCOSSMHNHSLPALSS 400
DB 355 SVIKESDHSTELDDITVTDYLSATKVSFDDTCLASE 392

RESULT 10
US-09-077-675A-2
Sequence 2, Application US/09077675A
Patent No. 6242199

GENERAL INFORMATION:
APPLICANT: Pal, Lee-Yuh
APPLICANT: Feighner, Scott C.
APPLICANT: Howard, Andrew D.
APPLICANT: Pong, Sheng-Shung
APPLICANT: Van Der Ploeg, Leonardus H.T.
TITLE OF INVENTION: RECEPTOR ASSAY
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000, 126 E. Lincoln Ave.
CITY: Rahway
STATE: NJ
COUNTRY: USA

ZIP: 07065-0900
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/077,675A
FILING DATE: 3-JUN-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Cocuzzo, Anna L.
REGISTRATION NUMBER: 42,452
REFERENCE/DOCKET NUMBER: 19590P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 732-594-1273
TELEFAX: 732-594-4720
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 302 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-077-675A-2

Query Match 18.4%; Score 402; DB 4; Length 302;
Best Local Similarity 33.6%; Pred. No. 3.2e-30;
Matches 95; Conservative 64; Mismatches 96; Indels 28; Gaps 7;

QY 68 LVLIHQAMKTPNTNYLFSLSVSDLLVL-LGMPLEVEMRNYPFLFGPYCFKALFE 127
DB 2 LVVSRREMTTNILVLSNAFSDLLIFLC-MPLDLFRLMQYRPNLGNLCKLPQVSE 60
QY 128 TVCFASILSTTVSVERYVAIHPRAKLOSTRRLRIIGIYWGFSVLSLPTNTHGT 187
DB 61 SCTAATVLTITLALSVREYFAICPLRAKVVTGKVKLVTLVIMAAVFCAGPIFVLGV 120
QY 188 KHFYFNGS-----LVPGSATCVIKPMWITVNTIQVTSFLFLPMTVTSYL 235
DB 121 E---HDNGIDPRTNCRATERFAVRSGLLTVM---VWV-----SSVFELFVCLITVL 167
QY 236 YVLMALRLKKDSLEADGNANIQRCKRSVNMFLVVLVFAICWAFHIDRLFFS-FV 294
DB 168 YSLIGRKLMMRRKGEAAVG-SLRDQNHKQTVKMLAVVFAFLICMLPFHVGRYLFPSKL 226
QY 295 EEMSESLAAVFNLVHYVSGYFFLLSSAVNPITYNLSRRQA 337
DB 227 EPGSVETIAQISQYCNLSFVFLYLSAINDILYNMSKKYRA 269

RESULT 11
US-09-077-675A-7
Sequence 7, Application US/09077675A
Patent No. 6242199

GENERAL INFORMATION:
APPLICANT: Pal, Lee-Yuh
APPLICANT: Feighner, Scott C.
APPLICANT: Howard, Andrew D.
APPLICANT: Pong, Sheng-Shung
APPLICANT: Van Der Ploeg, Leonardus H.T.
TITLE OF INVENTION: RECEPTOR ASSAY
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000, 126 E. Lincoln Ave.
CITY: Rahway
STATE: NJ


```

; ADDRESSEE: Jacobson, Price, Holman & Stern, PLLC
; STREET: 400 Seventh Street
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/472,880
; FILING DATE: 28-Dec-1999
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR 9723204
; FILING DATE: 17-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Player, William E.
; REGISTRATION NUMBER: 31,049
;
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 410 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-472-880-2

```

```

Query Match      18.2%; Score 398; DB 4; Length 410;
Best Local Similarity 29.2%; Pred. No. 1.1e-29;
Matches 106; Conservative 77; Mismatches 118; Indels 62; Gaps 13;

```

```

QY 49 SVYVPTFVGVGNVLCVLIQHOMAKT-PTNYVLFSLAVSDLVLLGMPLEYVE-M 106
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 36 TALYALIMWAGAGNMLSVHVLKARAGRGRHNLSTLALGLLLLVGVVELYSFV 95
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 107 WRVYPTFLGPGVCG---YFKALFETVOCFASILSTTVSVRVAAILHPRAKIQTSPRRA 163
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 96 WFRYPMVFGDLGRGYF---VHELCAIVATVLSVAGLSAERCLAVCOPLRARSLLPRRT 132
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 164 LRLIGVWGSFVLSLPTNSIHGKIFHY-FPNGSLVPGSATCTVIRKPMIYNFIQVTSF 222
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 153 RMVLVSMASGLGLAPMAIYMGQKHELEADEPPEPASVNCYIVLSRTALQVFIQVNL 212
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 223 LFLVLPRTVLSVL-----YLMAL-----RLKKDSLEADEG----- 234
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 213 VSFVLPALTAFLNGVTVSHLLACQVPSSTPGSSSTPSRLK-----LISEEGLSPFIV 268
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 255 -NANIO-----RPCRKSVMKMLFVLVLFVFAICWAPFHIDLFPSFV--EE 296
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 269 KRFITGGGOVSLVRHKDVRIRSLQNSV-OVLAIYVMVYICWLPCHARLMTGYPPDA 327
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 297 WSESLAIVNLVHVNSGVFFYLSAVNPFIYINLSRFRQAFQNVISFHKQWHSQHPQ 356
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 328 WTPDLYNFEHYFYWMTLTFLYVSSAVTPLYINAVSSFRKLFLDAVSSSLGGEHHPM--KR 385
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 357 LPP 359
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 386 LPP 388

```

RESULT 14
US-07-629-1041-3

; Sequence 3, Application US/076291041
; Patent No. 5288621
; GENERAL INFORMATION:
; APPLICANT: Gershenorn, Marvin C
; APPLICANT: Straub, Richard E
; TITLE OF INVENTION: PITUITARY TRH RECEPTOR
; NUMBER OF SEQUENCES: 3

```

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Yahwak & Associates
; STREET: 25 Skytop Drive
; CITY: Tumbull
; STATE: Connecticut
; COUNTRY: USA
; ZIP: 06611
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Microsoft Word 3.0
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/629,1041
; FILING DATE: 19901214
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: George M. Yahwak
; REGISTRATION NUMBER: 26,824
; REFERENCE/DOCKET NUMBER: CRF D - 995
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (203)268-1951
; TELEFAX: (203)268-1951
;
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 393 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-07-629-1041-3

```

```

Query Match      18.2%; Score 397; DB 1; Length 393;
Best Local Similarity 28.6%; Pred. No. 1.3e-29;
Matches 110; Conservative 72; Mismatches 104; Indels 98; Gaps 15;

```

```

QY 18 LEDPFOKHNSTK-----EY-----LAFLCGPRSRSHFPLPVSVYVPFVGVY 61
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1 MENDTVSENMQTELOPQAAVALEFQVTVLLVITIGC-----LGIV 41
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 62 GNVVLCVLIQHOMAKTPTNYVLFSLAVSDLVLL-LGML---EYEMKRNTPFLGCV 117
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 42 GNMVVLVWRTKHMRTPTNCYLVSLAVADLVVAAGLPNITDSYGSV---YGYV 96
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 118 GCVKALFETVOCFASILSTTVSVRVAAILHPRAKIQTSPRRLRLIGVWGSFV 177
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 97 GCICITYLYQVGINASSCSITTAFTIERIYAIQPIKAOFLCTSRRAKIIIFVMAFTSI 156
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 178 SLPTNSIHGKIFHYFPNGSLVPGSATC-----TVIKPMIYNFIQVTSFLLPMTV 231
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 157 CMLMFFLLDINISYTKNAVYV---SCGYKISRNYVSPYIMDFV-----FYVPMIL 206
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 232 ISVLYTYMA-----LRLKKDSLEADEG--NANIQRPC-----RKSVMK 268
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 207 ATVLXGFIARILFLNPIDSPDKENSKMKNDSIHOKNMLNATNCFNSTVSSRKQYTK 266
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 269 MLFVLVLFVFAICWAPFH---IDRLFPS--FVEEMSESLAIVNLVHVNSGVFFYLSAVN 323
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 267 MLAVVILFALMLMPTKTLVYVNSFLSSPQENM-----FLDFICICILYLSAIN 316
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 324 PIYINLSRFRQAFQNVISSFHK 347
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 317 PVYINLMSQKFRAPFRKLCNCKOK 340

```

RESULT 15
US-08-288-663A-15

; Sequence 15, Application US/08288663A
; Patent No. 5879896
; GENERAL INFORMATION:
; APPLICANT: HINUMA, Shuji
; APPLICANT: HOSOYA, Masaki
; APPLICANT: ONDA, Haruo

...

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2002, 14:17:17 ; Search time 17.05 Seconds

(without alignments)
892.428 Million cell updates/sec

Title: US-09-609-146-4

Perfect score: 2185

Sequence: 1 MSGMEKLQNASWYQOKLED.....ALSSQMSRTNYQSFHFNKT 415

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

100059

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	ID	Description
1	489.5	22.4	424	1	NTR1_RAT
2	486.5	22.3	424	1	NTR1_MOUSE
3	476	21.8	412	1	GP38_HUMAN
4	471	21.6	418	1	NTR1_HUMAN
5	448	20.5	366	1	GHSR_PIG
6	444	20.3	366	1	GHSR_HUMAN
7	436.5	20.0	364	1	GHSR_RAT
8	410	18.8	416	1	NTR2_RAT
9	408	18.7	395	1	TRFR_CHICK
10	405.5	18.6	398	1	TRFR_HUMAN
11	398	18.2	410	1	NTR2_HUMAN
12	397	18.2	393	1	TRFR_MOUSE
13	394.5	18.1	398	1	TRFR_SHEEP
14	393	18.0	412	1	TRFR_RAT
15	389.5	17.8	398	1	TRFR_BOVIN
16	387	17.7	417	1	NTR2_MOUSE
17	372	17.0	402	1	NK2R_CAVPO
18	365	16.7	369	1	SSR2_HUMAN
19	363	16.6	453	1	GP39_HUMAN
20	359.5	16.5	369	1	SSR2_PIG
21	359	16.4	368	1	SSR2_BOVIN
22	358	16.4	369	1	SSR2_RAT
23	357.5	16.4	399	1	BR31_CAVPO
24	357	16.3	398	1	OPRM_MOUSE
25	357	16.3	401	1	OPRM_BOVIN
26	355	16.2	400	1	OPRM_HUMAN
27	352.5	16.1	380	1	OPRM_RAT
28	352	16.1	401	1	OPRM_PIG
29	351.5	16.1	385	1	NK3R_MOUSE
30	351	16.1	384	1	SSR4_RAT
31	351	16.1	460	1	OX2R_RAT
32	350.5	16.0	380	1	OPRK_CAVPO
33	350.5	16.0	380	1	OPRK_HUMAN

34	350	16.0	384	1	SSR4_MOUSE	P49660 mus musculus
35	350	16.0	408	1	NK1R_RANCA	O98982 rana catesb
36	349	16.0	369	1	SSR2_MOUSE	P30875 mus musculus
37	348.5	15.9	380	1	OPRK_MOUSE	P33534 mus musculus
38	348.5	15.9	564	1	SH1L_DROME	P20905 drosophila
39	347.5	15.9	391	1	SSR1_HUMAN	P30872 homo sapien
40	347.5	15.9	391	1	SSR1_MOUSE	P30873 mus musculus
41	347.5	15.9	391	1	SSR1_RAT	P28646 ratius norv
42	346.5	15.9	363	1	SSR5_RAT	P30938 ratius norv
43	346	15.8	444	1	OX2R_CANFA	O91up7 cantis fam11
44	345.5	15.8	452	1	NK3R_RAT	P16177 ratius norv
45	344.5	15.8	398	1	OPRM_RAT	P33535 ratius norv

ALIGNMENTS

RESULT ID	1	STANDARD:	PRT:	424 AA.
AC	NTR1_RAT			
AC	P20789:			
DT	01-FEB-1991 (Rel. 17, Created)			
DT	01-FEB-1991 (Rel. 17, Last sequence update)			
DT	15-JUL-1998 (Rel. 36, Last annotation update)			
DE	NEUROTENSIN RECEPTOR TYPE 1 (NT-R-1) (HIGH-AFFINITY LEVOCABASTINE-SENSITIVE NEUROTENSIN RECEPTOR) (NTRH).			
GN	NTR1.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_Taxid=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	TISSUE=Brain;			
RX	MEDLINE=90297956; PubMed=1694443;			
RA	Tanaka K., Masu M., Nakanishi S.;			
RT	Structure and functional expression of the cloned rat neurotensin receptor.;			
RL	Neuron 4:847-854(1990).			
CC	-1- FUNCTION: RECEPTOR FOR THE TRIDECAPEPTIDE NEUROTENSIN. IT IS ASSOCIATED WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.			
CC	-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.			
CC	-1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS. HIGHEST TO TACHYKININS RECEPTORS.			
DR	P1R, JH0164, JH0164.			
DR	GCRD; GCR_0219; -			
DR	InterPro: IPR000276; GPCR_Rhodopsn.			
DR	Pfam: PF00001; 7tm.1; 1.			
DR	PRINTS: PR00237; GPCRHOPOPSN.			
DR	PROSITE: PS00237; G-PROTEIN_RECEP_FL_1; 1.			
DR	PROSITE: PS00262; G-PROTEIN_RECEP_FL_2; 1.			
KW	G-protein coupled receptor; Transmembrane; Glycoprotein; Phosphorylation; Lipoprotein; Palmitate.			
FT	DOMAIN 1 64			
FT	TRANSMEM 65 87			
FT	DOMAIN 88 96			
FT	TRANSMEM 97 121			
FT	DOMAIN 122 143			
FT	TRANSMEM 144 165			
FT	DOMAIN 166 188			
FT	TRANSMEM 189 210			
FT	DOMAIN 211 235			
FT	TRANSMEM 236 260			
FT	DOMAIN 261 308			
FT	TRANSMEM 309 330			
FT	DOMAIN 331 348			
FT	TRANSMEM 349 372			
FT	DOMAIN 373 424			
FT	CARBOHYD 4			
FT	CARBOHYD 38			
FT	CARBOHYD 42			
FT	CARBOHYD 42			
FT	DISULFID 142 225			

BY SIMILARITY.

FT	LIPID	388	388	PALMITATE (POTENTIAL).
SO	SEQUENCE	424 AA;	47054 MW;	ASC2FEAR8D9BCD3 CRC64:
Query Match		22.4%;	Score 489.5;	DB 1; Length 424;
Best Local Similarity		32.3%;	Pred. No. 1.6e-23;	
Matches 107;	Conservative	75;	Mismatches 114;	Indels 35; Gaps 9;
OY	48	VSVVVVPIFEVVCVGNVLCVLIIDH---	QANKTFTNYLFLSLAVSDLLVLLGMPLEVY	104
Db	67	VTATLALFEVVGTVGVSNSYATFTLAKKSLDSQSTVHNHGLSLASDLILLAPRVELY	126	
OY	105	E-MMRNRYEFLLEGPGVC---YFKTALFEYCEFASLSTTVSYERYVAIILHPRAKLOSTR	160	
Db	127	NFIWVHHPMAGDACCGRGYE---LRDACTVATALNVA SLSEYERLALCHPRKATIMSR	183	
OY	161	RRAIRLILGWGFSVLSFSPNTSHIGLIFHFHPNGSLVPGSATCIVIKPMIYNFIIQVT	220	
Db	184	SRTKKEFIAIWMASLLALIPMLFTGLGNR---SGDGTHPGGVLCPTVIDATAVKVIQVN	241	
OY	221	SFLFLYLPMTVIVSVYYUMLALRLKDKDSLAEDEG-----NANIORPCRK	264	
Db	242	TFMSGLFLMFLVLSLINTYAIYANKLYTMVIOALAEQGRVCYVGHNGLEHSTFNKTFIE-PGRV	300	
OY	265	SVNK---MLFVLVLVFAICWAPFHIIDLRFSEV--EEMSESIAAVFNLVHVVCVFVFL	318	
Db	301	QALRHGVLVLRAVVYAFVAVCWPLRYHVRRILMEFCYIDEDQWTFELFDGYHYFVWLNALFEVY	360	
OY	319	SSAVNPPIIYNLLSRFOAAFOVVISFRKQW	349	
Db	361	SSAIPPIILYNLSANFROVFLSTLACICPGW	391	

RESULT	2	
ID	NTRL_MOUSE	
AC	088319	STANDARD: PRT: 424 AA.
DT	20-AUG-2001 (Rel. 40, Created)	
DT	20-AUG-2001 (Rel. 40, Last sequence update)	
DT	20-AUG-2001 (Rel. 40, Last annotation update)	
DE	NEUROTENSIN RECEPTOR TYPE 1 (NT-R-1).	
GN	NTSL OR NTSR.	
OS	Mus musculus (Mouse).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
OX	NCBI_Taxid=10090;	
PN	[1]	
RP	SEQUENCE FROM N.A.	
RA	TISSUE-Brain;	
RC	Snider J., Sano H., Ohla M.;	
RL	"Neurotensin receptor type 1 ";	
RL	Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.	
CC	-1- FUNCTION: RECEPTOR FOR THE TRIDECAPEPTIDE NEUROTENSIN. IT IS ASSOCIATED WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.	
CC	-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.	
CC	-1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS. HIGHEST TO TACHIKININS RECEPTORS.	
CC	-----	
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	
CC	the European Bioinformatics Institute. There are no restrictions on its	
CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (See http://www.isb-sib.ch/announcement/	
CC	or send an email to license@sib-sib.ch).	
CC	-----	
DR	EMBL; AB017027; BAA33013.1; -	
DR	MGI; MGI:97386; Ntsr.	
DR	InterPro; IPR000276; GPCR_Rhodpsn.	
DR	Pfam; PF00001; 7tm.1; 1.	
DR	PRINTS; PR00237; GPCRHHODPSN.	
DR	PROSITE; PS00237; G_PROTEIN_RECPT_FL_1; 1.	

Query Match	Score	DB 1	Length
Best Local Similarity	30.9%	Pred. No. 2.4e-23	424
Matches 116: Conservative	79:	Mismatches 126:	Indels 53: Gaps 11:

	Query Match	22.3%	Score 486.5;	DB 1:	Length 424;	
	Best Local Similarity	90.9%	Pred.	No. 2.4e-23;		
	Matches	116;	Conservative	79;	Mismatches 120; Indels 53; Gaps 11.	
QY	7 LONASMIYOOKLEDPPOKHLNSTEERYLAFLCGPRRSHEFLLPVSVVVYPPIFFVGVIGNVLYL	66				
Dd	LSNGSGNSESESLBPNMSLDVNTDIYSKYLV-----VTAYVALAFVVGTVGNSTV	84				
QY	67 CLVIYLQH--QAAMPTNYLFELSLAVSDILLVLCLMPLVEY-MRRNPPELPYGCC---	119				
Dd	AFTLARKKSLQSLOSTVHYHGLSIALSIDILLMLAMPVELYNFIWVHHPMAGDAGCARGY	144				
QY	120 YEKTALEPFCVASLTITTVSVEEVAIVILHPFAKRLQSTRRALRIIGIWGFSVTLST	179				
Dd	YF---LRACCTATPALNANSLVSVERVLIICHFEKKTLMSSRKRTKFSAIMLASALLAV	201				
QY	180 PNTSHGIRKIFHFPPGSU---VPSSACTGTVICPKMYINTEIIQVTSFLFYLBMTYISVLY	236				
Dd	202 PMLEFMGIQ-----NRSMADGGHPGSLVCETPYDTATVKVLIQVNFMSSLFPMILIISTLN	256				
QY	237 YLMALRLK-----KDKSLEADEGCNNANIQRCKRCSYVK---MLEFVLYL	275				
Dd	257 TVIANKLTLVMYHOAAEOGRGYCVTGTNHNSLEHSTFNMSIE-PGRAOQLAHGVLVRAVVI	315				
QY	276 VEAICMAPPHIDRLFESV--EEMSESILAAYNLVHVYSCGFYISSAENVPIIYNLNSRR	333				
Dd	316 AFVCMGLEPHYHRRLMEFCIISDEOMTFLEDFYHYHYMLTNALFEVSSAINPIDILNLYSAN	375				
QY	334 FOAAFQONYVISPFHKQM	349				
Dd	376 FROYFLSTLACICPGM	391				
RESULT	3					
GP38_HUMAN	ID	GP38_HUMAN	STANDARD;	PRT;	412 AA.	
AC	O43193;					
DT	15-DEC-1998 (Rel. 37, Created)					
DT	15-DEC-1998 (Rel. 37, Last sequence update)					
DT	20-AUG-2001 (Rel. 40, Last annotation update)					
DE	PUTATIVE G PROTEIN-COUPLLED RECEPTOR GP38.					
GN	GP38.					
OS	Homo sapiens (Human).					
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.					
X	NCBI Taxid=9606;					


```

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98110578; PubMed=9441746;
RA McKee K.K., Tan C.P., Palyna O.C., Liu J., Feighner S.D.,
Hreniuk D.L., Smith R.G., Howard A.D., van der Ploeg L.H.T.;
RT "Cloning and characterization of two human G protein-coupled receptor
genes (GPR38 and GPR39) related to the growth hormone secretagogue
RT and neurotensin receptors."
RL Genomics 46:426-434(1997).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: EXPRESSED ONLY IN THYROID, STOMACH, AND BONE
CC MARROW.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF034632; AAC26081.1; -.
DR GCRDB; GCR_2494; -.
DR MIM; 602885; -.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm.1; 2.
DR PRINTS: PR00237; GPCR_RHODOPSN.
DR PROSITE: PS00237; G_PROTEIN_REC_P1_1; 1.
DR PROSITE: PS0262; G_PROTEIN_REC_P1_2; 1.
KM G-protein coupled receptor; Transmembrane; Glycoprotein.
FT DOMAIN 1 35 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 36 56 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 57 74 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 75 94 2 (POTENTIAL).
FT DOMAIN 95 112 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 113 134 3 (POTENTIAL).
FT DOMAIN 135 157 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 158 178 4 (POTENTIAL).
FT DOMAIN 179 246 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 247 270 5 (POTENTIAL).
FT DOMAIN 271 298 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 299 320 6 (POTENTIAL).
FT DOMAIN 321 334 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 335 358 7 (POTENTIAL).
FT DOMAIN 359 412 CYTOPLASMIC (POTENTIAL).
FT DISULFID 111 235 BY SIMILARITY.
FT CARBOHYD 6 6 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 192 192 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 412 AA; 45344 MW; C13FF6165012DEF3 CRC64;

```

```

Query Match 21.8%; Score 476; DB 1; Length 412;
Best Local Similarity 31.8%; Pred. No. 1e-22;
Matches 115; Conservative 81; Mismatches 106; Indels 60; Gaps 11;

```

```

QY 37 CQPRSHFLPRVSVYVFFVGVGNVLCVLIIDHQAQKPTNYFLSLAVSLVLL 96
DB 1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
30 CQPRPALGALVPTAVACLCLEFVVGSGNVTVMLIGRYRMRTTNLYLSMAVSDLLI 88
QY 97 LQMPLEVEEMRNVPFLFCVPCYCFKTALEFVVCFAISITTVSERYVAIIHPRAKL 156
DB 89 LQLPDPLVLTMSRRVFPVPLCRSLVLYGEGCTVATLHMTALSERIALCRPLARAY 148
QY 157 QSTRRRALILCIWGFVSFLSPNTSHIGIKFHFPGNSLVPG-----SATCTVIR 208
DB 149 LVTRRVRLIIVLAVALLSAGPELVGVE--ODPGISVPGGLGTARIASSPLASSP 206
QY 209 PMMI-----INFIQVTSLEFYLPMVIVS 234
DB 207 PLMLSRAPPSPSPSGEETAALALSRRCRSPDAGLRVLMWTTAAVF-LPFLCLSI 265
QY 235 LVYLMALRLKKDKSLKADGEGNANIORPKRSYKMLFVLVLAICWAFPHIDRLFEFSV 294

```

```

DB 266 LVGLIGRELMSRRPLRGPAASGRGHRQTV-RVLLVVLAFIICWLPFHGRITVIYNT 324
QY 295 EEMSESL--AAVENIVHVSVGFYLLSSAVNPITVLLSRRRO-AAFOVVIS-----SF 345
DB 325 ED-SRMVFSYVENIVAL--QLFYLSASINPLVNLISKRYAAAFKLLARKSRPGE 380
QY 346 HK 347
DB 381 HR 382
RESULT 4
ID NTBL_HUMAN STANDARD; PRT; 418 AA.
AC P30989;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE NEUROTENSIN RECEPTOR TYPE 1 (NT-R-1) (HIGH-AFFINITY LEVOCABASTINE-
DE INSNSITIVE NEUROTENSIN RECEPTOR) (NTRH).
OS NTR1 OR NTRR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93154505; PubMed=8381365;
RA Vita N., Laurent P., Lefort S., Chalon P., Dumont X., Kaghad M.,
Gully D., Le Fur G., Ferrara P., Caput D.;
RT "Cloning and expression of a complementary DNA encoding a high
RT affinity human neurotensin receptor."
RL FEBS Lett. 317:139-142(1993).
CC -1- FUNCTION: RECEPTOR FOR THE TRIDECAPEPTIDE NEUROTENSIN. IT IS
CC ASSOCIATED WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-
CC CALCIN SECOND MESSENGER SYSTEM.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC HIGHEST TO TACHIKININS RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X70070; CAA49675.1; -.
DR PIR; S29506; S29506.
DR GCRDB; GCR_0577; -.
DR MIM; 162651; -.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm.1; 1.
DR PROSITE: PS00237; G_PROTEIN_REC_P1_1; 1.
DR PROSITE: PS0262; G_PROTEIN_REC_P1_2; 1.
KM G-protein coupled receptor; Transmembrane; Glycoprotein;
KW Phosphorylation; Lipoprotein; Palmitate.
FT DOMAIN 1 63 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 64 86 1 (POTENTIAL).
FT DOMAIN 87 95 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 96 120 2 (POTENTIAL).
FT DOMAIN 121 142 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 143 164 3 (POTENTIAL).
FT DOMAIN 165 187 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 188 209 4 (POTENTIAL).
FT DOMAIN 210 234 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 235 239 5 (POTENTIAL).
FT DOMAIN 260 303 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 304 325 6 (POTENTIAL).
FT DOMAIN 326 343 EXTRACELLULAR (POTENTIAL).

```

```

FT TRANSMEM 344 367 7 (POTENTIAL).
FT DOMAIN 368 418 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 37 4 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 37 37 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 41 41 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 141 224 BY SIMILARITY.
FT LIPID 383 383 PALMITATE (POTENTIAL).
SQ SEQUENCE 418 AA: 46288 MW: 8881ECC2BE6390 CRC64;

Query Match 21.6%; Score 471; DB 1; Length 418;
Best Local Similarity 31.0%; Pred. No. 2.1e-22;
Matches 120; Conservative 72; Mismatches 121; Indels 74; Gaps 14;

OY 20 DPQKHLNSTEYL-----AFICGPRRS-----HFLPVSVVYVPEV 58
DB 17 DPQKHLNSTEYL-----AFICGPRRS-----HFLPVSVVYVPEV 76
OY 59 GVIGNVLYCLVLIQH---QAMKTPNYLYFLSLAVSDLLVLLGMPLEVE-MMRNPELF 114
DB 77 GYGNVTYATFLARKKSLQSLQSTVHYHGLSLASDLTLTLLAMPVELYVFWHHWAF 136
OY 115 GPVGC---YKTLAFETVCASLSTTTSVERVVALHPFRAKLOSTRRLRLIGI 171
DB 137 GDAGCKROYF---LRDAGCTATALNVAASLSVEHYLAICHPFAKTLMSRSRTEKFTSAIW 193
OY 172 GFGLVSLP-----NTSIGHIKFHPNGSLV--PGSATCTYIKPMIYNFIQVTS 221
DB 194 LASALTLVPLFTMGEDNKRADQH-----AGGLVCTPTIHTIV-----KVIOVNT 241
OY 222 FLELYLPMYIVSYLYLMLRLKDKLSLEADEN-----ANIOPCRK 264
DB 242 FMSFIPMYIVSYLYLMLRLKDKLSLEADEN-----ANIOPCRK 300
OY 265 SVKMLFVLVLAICAPRHIDLPFSEV--EEMSESLAVNLVHVSCVFYISSAV 322
DB 301 GV--RVLRAVVAIVVCMPLRVHRLMFCYISDQWTPFLDYHFYFVMTNALFVYSTI 359
OY 323 NPITYNLRSRFOAFONVISFHKW 349
DB 360 NPILYNVSNFRHIFLATIACLPW 386

RESULT 5
GHSR_PIG STANDARD: PRT: 366 AA.
AC 095254: 095255:
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE GROWTH HORMONE SECRETAGOGUE RECEPTOR TYPE 1 (GHS-R) (GH-RELEASING
DE PEPTIDE RECEPTOR (GHRP).
GN GHSR
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_Taxid=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FORKSHIRE; TISSUE=Pituitary;
RX MEDLINE=96337998; PubMed=8688086;
RA Howard A.D., Feighner S.D., Cully D.F., Arena J.P.,
RA Liberator P.A., Rosenblum C.I., Hamelin M., Hreniuk D.L.,
RA Palyha O.C., Anderson J., Paress S., Diaz C., Chou M., Liu K.K.,
RA McKee K.K., Pong S.-S., Chaung L.-Y., Elbrecht A., Daskewicz M.,
RA Heavens R., Rigby M., Strimachsinghji D.J.S., Dean D.C., Mellilo D.G.,
RA Patchett J.M., Nargund R., Griffin P.R., Demartino J.A., Gupta S.K.,
RA Schaeffer J.A., Smith R.G., van der Ploeg L.H.T.;
RA "A receptor in pituitary and hypothalamus that functions in growth
RT hormone release."
RL Science 273:974-977(1996).
CC -!- FUNCTION: RECEPTOR FOR GROWTH HORMONE RELEASING PEPTIDES (GHRP) AS
CC WELL AS NON-PEPTIDE, LOW MOLECULAR WEIGHT SECRETAGOGUES (E.G. L-
```

```

CC 692,429, MK-0677). THIS RECEPTOR IS COUPLED TO G-ALPHA-11
CC PROTEIN. BINDING POTENCY FOR THE SECRETAGOGUES IS IN THE ORDER:
CC MK-0677 > GHRP-2 > GHRP-6.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1A (SHOWN HERE) AND 1B; ARE
CC PRODUCED BY ALTERNATIVE SPLICING. ISOFORM 1B APPEARS NOT TO BIND
CC SECRETAGOGUES.
CC -!- TISSUE SPECIFICITY: PITUITARY AND HYPOTHALAMUS.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed, usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U60178; AAC48630.1; -
DR EMBL: U60180; AAC48631.1; -
DR GCRDB: GCR_1532; -
DR GCRDB: GCR_1538; -
DR InterPro: IPR003905; GHS1_receptor.
DR InterPro: IPR002276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PR00237; GPCRHHODOPSN.
DR PROSITE: PS00237; G_PROTEIN_RECP_FL_1; 1.
DR PROSITE: PS0262; G_PROTEIN_RECP_FL_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
KW Alternative splicing.
FT DOMAIN 1 40
FT TRANSMEM 41 66
FT DOMAIN 67 72
FT TRANSMEM 73 96
FT DOMAIN 97 117
FT TRANSMEM 118 139
FT DOMAIN 140 162
FT TRANSMEM 163 183
FT DOMAIN 184 211
FT TRANSMEM 212 235
FT TRANSMEM 236 263
FT DOMAIN 264 285
FT TRANSMEM 286 302
FT TRANSMEM 303 326
FT DOMAIN 327 366
FT DISULFID 116 198
FT CARBOHYD 13 13
FT CARBOHYD 27 27
FT VARSPLIC 266 289
FT VARSPLIC 290 366
SQ SEQUENCE 366 AA: 41194 MW: 2C850B3BEF61B701C CRC64;

Query Match 20.5%; Score 448; DB 1; Length 366;
Best Local Similarity 34.3%; Pred. No. 4.8e-21;
Matches 104; Conservative 68; Mismatches 103; Indels 28; Gaps 7;

OY 48 VSVYVPIPVGVYIGNVLCVLIHQAMKTPNYLYFLSLAVSDLLVLLGMPLEVEYEM 107
DB 46 VYVYVPIPVGVYIGNVLCVLIHQAMKTPNYLYFLSLAVSDLLVLLGMPLEVEYEM 104
OY 108 RNPPLFGPGVGCYKFLAFETVCASLSTTTSVERVVALHPFRAKLOSTRRLRLI 167
DB 105 QYPMNLGNLCLKLQFVSSCTYATVLTITLAVERTFAICPLKAKVYVTKGRKLVY 164
OY 168 GYVGCYSVLSFSLNNTSIGHIKFHPNGS-----LVGSATCTYIKPMIYN 215
DB 165 LVIMVAFAFCAGPIFLVAGV---HNGTDPDRPTNCRATEFAVRGGLTYM--VWV--- 216
OY 216 IIOVTSFELYLPMYIVSYLYLMLRLKDKLSLEADENANIOPCRKRSVKNMFLVLY 275
DB 217 -----SSVEFLFVFLVLYSLIGRKLWRRKGEAAVG--SSIRDQNHQOTVKNLAVVF 270
```


RL Mol. Endocrinol. 11:415-423(1997).
 RN (2)
 RP SEQUENCE OF 1-240 FROM N.A.
 RC STRAIN-WISTAR: TISSUE-Pituitary;
 RX MEDLINE-98100386; PubMed-9437732;
 RA Yokote R., Sato M., Matsubara S., Ohye H., Nimi M., Murao K.,
 RA Takahara J.;
 RT "Molecular cloning and gene expression of growth hormone-releasing
 RT peptide receptor in rat tissues.";
 RT Peptides 19:15-20(1998)
 CC -1- FUNCTION: RECEPTOR FOR GROWTH HORMONE RELEASING PEPTIDES (GHRP) AS
 CC WELL AS NON-PEPTIDE, LOW MOLECULAR WEIGHT SECRETAGOGUES (E.G. L-
 CC 692,429, MK-0677). THIS RECEPTOR IS COUPLED TO G-ALPHA-11
 CC PROTEIN (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U94321; AAC53156.1; -
 DR EMBL: AB001982; BAA2177.1; ALT_INT.
 DR GCRDB: GCR_1383; -
 DR InterPro: IPR003905; GHS1_receptor.
 DR InterPro: IPR002276; GPCR_Rhodopsn.
 DR Pfam: PF00001; 7tm_1; 1.
 DR PRINTS: PR00237; GPCRHHODPSN.
 DR PROSITE: PS00237; G-PROTEIN_RECP_F1_1; 1.
 DR PROSITE: PS50262; G-PROTEIN_RECP_F1_2; 1.
 KM G-protein coupled receptor; Transmembrane; Glycoprotein.
 FT DOMAIN 1 40
 FT TRANSSEM 41 66
 FT DOMAIN 67 72
 FT TRANSSEM 73 96
 FT DOMAIN 97 117
 FT TRANSSEM 118 139
 FT DOMAIN 140 162
 FT TRANSSEM 163 183
 FT DOMAIN 184 211
 FT TRANSSEM 212 235
 FT DOMAIN 236 263
 FT TRANSSEM 264 285
 FT DOMAIN 286 302
 FT TRANSSEM 303 326
 FT DOMAIN 327 364
 FT DISULFID 115 197
 FT CARBOHYD 13 13
 FT CARBOHYD 26 26
 SQ SEQUENCE 364 AA; 40963 MW; DCBF539BE061EE9 CRC64;

Query Match 20.0%; Score 436.5; DB 1; Length 364;
 Best Local Similarity 34.0%; Pred. No. 2.4e-20;
 Matches 103; Conservative 68; Mismatches 103; Indels 29; Gaps 8;

Db 216 -----SSVFFFLPVECTLVLYSLGRKLMRRG-DAAVG-ASLRDQHKTQVMAVVVF 268
 QY 276 VFAICAPPHIDLFPS-FVEEMSESLAAVFNLYHYSGVFFLLSVAVNPITINLSRRF 334
 Db 269 AFLICWLPFHVGRYLFKSPSPGLSTIOISYCNVSEVFLYLSAINDILVNMISKY 328
 QY 335 QAA 337
 Db 329 RVA 331
 RESULT 8
 NR2_RAT
 ID NR2_RAT STANDARD; PRT; 416 AA.
 AC Q63384;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE NEUROTENSIN RECEPTOR TYPE 2 (NT-R-2) (HIGH-AFFINITY LEVOCABASTINE-
 DE SENSITIVE NEUROTENSIN RECEPTOR).
 GN NR2 OR NR2.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_Taxid=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Hypothalamus;
 RX MEDLINE-96228041; PubMed-8647296;
 RA Chalon P., Vita N., Kaghed M., Guillemont M., Bonin J.,
 RA Delpech B., Le Fur G., Ferrara P., Caput D.;
 RT "Molecular cloning of a levocabastine-sensitive neurotensin binding
 RT site.";
 RL FEBS Lett. 386:91-94(1996).
 CC -1- FUNCTION: RECEPTOR FOR THE TRIDECAPEPTIDE NEUROTENSIN. IT IS
 CC ASSOCIATED WITH G PROTEINS, THAT ACTIVATE A PHOSPHATIDYLINOSITOL-
 CC CALCULIN SECOND MESSENGER SYSTEM.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- TISSUE SPECIFICITY: ABUNDANT IN CORTEX AND HYPOTHALAMUS, AND LOWER
 CC LEVELS SEEN IN THE HEART AND INTESTINE.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED MAXIMALLY IN 7-DAY-OLD BRAIN AND
 CC EXPRESSION DECREASES PROGRESSIVELY UNTIL ADULTHOOD (35-DAY-OLD
 CC BRAIN).
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC HIGHEST TO TACHYKININS RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X97121; CA65787.1; -
 DR GCRDB: GCR_1461; -
 DR InterPro: IPR002276; GPCR_Rhodopsn.
 DR Pfam: PF00001; 7tm_1; 1.
 DR PROSITE: PS00237; G-PROTEIN_RECP_F1_1; 1.
 DR PROSITE: PS50262; G-PROTEIN_RECP_F1_2; 1.
 KM G-protein coupled receptor; Transmembrane; Lipoprotein; Palmitate.
 FT DOMAIN 1 32
 FT TRANSSEM 33 55
 FT DOMAIN 56 64
 FT TRANSSEM 65 87
 FT DOMAIN 88 109
 FT TRANSSEM 110 131
 FT DOMAIN 132 154
 FT TRANSSEM 155 176
 FT DOMAIN 177 216
 FT TRANSSEM 217 237
 FT DOMAIN 238 297
 CC CYTOPLASMIC (POTENTIAL).

AC 095665;
 DT 20-AUG-2001 (Rel. 40, Created)
 DT 20-AUG-2001 (Rel. 40, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE NEUROKININ RECEPTOR TYPE 2 (NT-R-2) (LEVOCABASTINE-SENSITIVE
 DE NEUROKININ RECEPTOR (NTR2 RECEPTOR).
 GN NTR2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=99066919; PubMed=9851594;
 RA Vito N., Oury-Donat F., Chalon P., Guillemot M., Kaghed M., Bachy A.,
 RA Thurensen S., Garcia S., Poinot-Chazel C., Casellas P., Keane P.,
 RA Le Fur G., Maffrand J.P., Shoubrite P., Caput D., Ferrara P.;
 RT "Neurotensin is an antagonist of the human neurotensin NT2 receptor
 RT expressed in Chinese hamster ovary cells.";
 RL Eur. J. Pharmacol. 360:265-272(1998).
 CC -1- FUNCTION: RECEPTOR FOR THE TRIDECAPEPTIDE NEUROTENSIN. IT IS
 CC ASSOCIATED WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-
 CC CALCIUM SECOND MESSENGER SYSTEM.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC HIGHEST TO TACHIKININS RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Y10148; CAA71233.1;
 DR MIM: 605538;
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR Pfam: PF00001; 7tm1.1;
 DR PRINTS: PR00237; GPCR_Rhodopsn.
 DR PROSITE: PS00237; G-PROTEIN_RECPEP_F1_1; 1.
 DR PROSITE: PS50262; G-PROTEIN_RECPEP_F1_2; 1.
 KW G-protein coupled receptor; Transmembrane; Lipoprotein; Palmitate.
 FT DOMAIN 1 32 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 33 55 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 56 64 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 65 87 2 (POTENTIAL).
 FT DOMAIN 88 109 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 110 131 3 (POTENTIAL).
 FT DOMAIN 132 154 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 155 176 4 (POTENTIAL).
 FT DOMAIN 177 217 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 218 237 5 (POTENTIAL).
 FT DOMAIN 238 297 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 298 318 6 (POTENTIAL).
 FT DOMAIN 319 337 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 338 358 7 (POTENTIAL).
 FT DOMAIN 359 410 CYTOPLASMIC (POTENTIAL).
 FT DISULFID 108 194 BY SIMILARITY.
 FT LIPID 377 377 PALMITATE (POTENTIAL).
 SQ SEQUENCE 410 AA: 45413 MW: 8C3ADA22BE15FD66 CRC64;

Query Match 18.2%; Score 398; DB 1; Length 410;
 Best Local Similarity 29.2%; Pred. No. 6e-18;
 Matches 106; Conservative 77; Mismatches 118; Indels 62; Gaps 13;

OY 49 SVYVPIFVGVYGVNVLVLQHQAMKT-PNIVYLSLAVSLDVLVLLGMPLEYE-M 106
 DB 36 TALVILIMLVAGNALSVHVKARAGRAGRHNVHLLAGLLLLLVGVVELYISFV 95
 OY 107 MNVYPLFPGVGC---YFKTALFETVCFASLITFTVSEYVALIHPRAKIQSTRRA 163

DB 96 WFIYPMVFGDLGGRGYF---VHELCAVATVLSVAGLSMRCLAVCOPLARASLLPRRT 152
 OY 164 LRLIGVWGEVSFLSPNTHSICRKH-PPNGLVPGSATCTVIRPMIYNTIOTYSF 222
 DB 153 RMLVALSMASLGLALPMVIMQCKHELETADEPEPASPVCVLSRPAQVFIQVMVL 212
 OY 223 LFLPLPMVYISLV-----YVLMAL-----RLKKDLSLEADEG----- 254
 DB 213 VSFVLPALAFPLANGVTVSHLLALCSQVPSTSPGSSTSRLE----LSEGLLSFIYW 268
 OY 255 -NANIQ-----RPCRKSVNKMFLVLVVAICAPPHIDRIFSFV--EE 296
 DB 269 KTFEIGGGVSLVRHKDVRIRISLQRSV-QVLAIVYVYICLPRHARLMCYPPDA 327
 OY 297 WSESLAAYNLVHVSGVFYSSAVNPITLYNLRSFQAFQNYISFFKQMSHQDPO 356
 DB 328 WTDPLNFYHYFMVNTLFYVSSAVTPLLYNVSSSFRKLFLEAVSSLCGEHHPM--KR 385
 OY 357 LPP 359
 DB 386 LPP 388

RESULT 12
 TRFR_MOUSE
 ID TRFR_MOUSE STANDARD; PRT; 393 AA.
 AC P21761;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-MAY-1991 (Rel. 18, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE THYROTROPIN-RELEASING HORMONE RECEPTOR (TRH-R) (THYROLIBERIN
 DE RECEPTOR).
 GN TRHR.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathia; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=pituitary;
 RX MEDLINE=91088548; PubMed=2175902;
 RA Straub R.E., Frech G.C., Joho R.H., Gershengorn M.C.;
 RT "Expression cloning of a cDNA encoding the mouse pituitary
 RT thyrotropin-releasing hormone receptor.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:9514-9518(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92381047; PubMed=1324930;
 RA Narayanan C.S., Fujimoto J., Geras-Raaka E., Gershengorn M.C.;
 RT "Regulation by thyrotropin-releasing hormone (TRH) of TRH receptor
 RT mRNA degradation in rat pituitary GH3 cells.";
 RL J. Biol. Chem. 267:17296-17303(1992).
 RN [3]
 RP SEQUENCE OF 332-393 FROM N.A.
 RC TISSUE=pituitary;
 RX MEDLINE=97013702; PubMed=9156522;
 RA Jones K.E., Brubaker J.H., Chin W.W.;
 RT "An alternative splice variant of the mouse TRH receptor mRNA is the
 RT major form expressed in the mouse pituitary gland.";
 RL J. Mol. Endocrinol. 16:197-204(1996).
 CC -1- FUNCTION: RECEPTOR FOR THYROTROPIN-RELEASING HORMONE. THIS
 CC RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE A
 CC PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@lsb-sib.ch.
 CC EMBL: M59811; AAA0480.1; -
 DR EMBL: M94384; AAA0437.1; -
 DR EMBL: L48936; AAA81559.1; -
 DR PIR: A39251; A39251.
 DR GCRDB: GCR_0099; -
 DR GCRDB: GCR_1613; -
 DR MGD: MGI:98824; Trhr.
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR Pfam: PF00001; 7tm_1; 1.
 DR PRINTS: PR00751; THYROLIBRIN.
 DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; 1.
 DR PROSITE: PS0262; G_PROTEIN_RECEP_F1_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein.
 FT DOMAIN 1 28
 FT TRANSMEM 29 51
 FT DOMAIN 52 61
 FT TRANSMEM 62 83
 FT TRANSMEM 84 99
 FT TRANSMEM 100 121
 FT TRANSMEM 122 144
 FT TRANSMEM 145 168
 FT TRANSMEM 169 193
 FT TRANSMEM 194 215
 FT TRANSMEM 216 266
 FT TRANSMEM 267 288
 FT TRANSMEM 289 296
 FT TRANSMEM 297 319
 FT TRANSMEM 320 393
 FT DOMAIN 3 3
 FT CARBOHYD 10 10
 FT CARBOHYD 3 3
 SQ SEQUENCE 393 AA; 44559 MW; 8739B75D1A0FCFC CRC64;

Query Match 18.28; Score 397; DB 1; Length 393;
 Best Local Similarity 28.68; Pred. No. 6.7e-18;
 Matches 110; Conservative 72; Mismatches 104; Indels 98; Gaps 15;

DB 18 LEAPFGKHNSTP-----EY-----LAFICGRSHFFLPVSVYVPIFVGVY 61
 1 MENDTVSENMOTELPOAAVALEYOVTVLLVYICG-----LGIV 41
 OY 62 GNVLVCLVLIHQAMKPTNYVLFSLAVSDLLVL--LGMPL--EYEMRNYPFLGPPV 117
 DB 42 GNMVVLVVMRTKHMRTPCNYLVSLAVADLVLAAGLNLNDSIGSM--YGVY 96
 OY 118 GCFKALPETYVCFASLITTVSVERYVALILPFRAKLOSTRRLRIIGIYWGESVLF 177
 DB 97 GCLCITYLYOYLGINASSCSITAFTEIERYIAICHPIKQFLCTFSRAKKIIFWAFPSIY 156
 OY 178 SLPTNSIHGIKFHYFPNGSLVPGSANC-----TVIKPMIYVETIQVTSFLFYLLPMY 231
 DB 157 CHLMFELDLINISTYKNNAVYV--SCGYKISKNYISPLIMDGV-----FYVPMIL 206
 OY 232 ISVLYYLMA-----LRLKDKSLEADG--NANTIORPC-----RKSVMK 268
 DB 207 ATVLGFIARILFLNPIDSPKENSMMKNDISHKNNLNLNTNRCFNSTVSSRKQVTK 266
 OY 269 MFLVLLVVAICAPRH---IDRLFFS--FVEEMSESLLAVNLVHVSVVFYLLSAVN 323
 DB 267 MALVVVILFALIMMPYRTLVVNSFLSSPQENW-----FLLECRICIYLNAIN 316
 OY 324 PIYVNLISRFQAAFOYVSSFKR 347
 DB 317 PVIYNLMSQKFAFRKLCNCKOK 340

RESULT 13
 TRFR_SHEEP STANDARD: PRT; 398 AA.
 AC Q28396;
 DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE THYROTROPIN-RELEASING HORMONE RECEPTOR (TRH-R) (THYROLIBRIN
 DE RECEPTOR).
 GN
 OS Ovis aries (Sheep).
 OC Eukaryota: Metazoa: Chordata: Vertebrata: Euteleostomi;
 OC Mammalia: Eutheria: Cetartiodactyla: Ruminantia: Pecora: Bovidae;
 OC Bovidae: Caprinae: Ovis.
 OC NCBI_TaxID=9940;
 RN
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Plutary;
 RX MEDLINE=97200775; PubMed=9048604;
 RA Bockmann J., Boeckers T.M., Winter C., Wittkowski W., Winterhoff H.,
 RA Deufel T., Kreutz M.R.;
 RT "Thyrotropin expression in hypophyseal pars tuberalis-specific cells
 RT is 3,5,3'-triiodothyronine, thyrotropin-releasing hormone, and pit-1
 RT independent.";
 RL Endocrinology 138:1019-1028(1997).
 CC -1- FUNCTION: RECEPTOR FOR THYROTROPIN-RELEASING HORMONE. THIS
 CC RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE A
 CC PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.lsb-sib.ch/announce/>
 CC or send an email to license@lsb-sib.ch).
 CC
 CC EMBL: X95285; CA64606.1; -
 DR GCRDB: GCR_1294; -
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR Pfam: PF00001; 7tm_1; 1.
 DR PRINTS: PR00751; THYROLIBRIN.
 DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; 1.
 DR PROSITE: PS0262; G_PROTEIN_RECEP_F1_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein.
 FT DOMAIN 1 28
 FT TRANSMEM 29 51
 FT DOMAIN 52 61
 FT TRANSMEM 62 83
 FT TRANSMEM 84 99
 FT TRANSMEM 100 121
 FT TRANSMEM 122 144
 FT TRANSMEM 145 168
 FT TRANSMEM 169 193
 FT TRANSMEM 194 215
 FT TRANSMEM 216 266
 FT TRANSMEM 267 288
 FT TRANSMEM 289 296
 FT TRANSMEM 297 319
 FT TRANSMEM 320 398
 FT CARBOHYD 3 3
 FT CARBOHYD 10 10
 SQ SEQUENCE 398 AA; 45088 MW; 375A311D3D2A61A CRC64;

Query Match 18.18; Score 394.5; DB 1; Length 398;
 Best Local Similarity 29.68; Pred. No. 9.6e-18;
 Matches 118; Conservative 74; Mismatches 129; Indels 77; Gaps 14;

OY 48 VSVVYPIFY-VGVIGNVCLVLIHQAMKPTNYVLFSLAVSDLLVL--LGMPL--E 102
 DB 27 VTLLVLIIGGLGIVGIMVVLVVMRTKHMRTPCNYLVSLAVADLVLAAGLNLNDS 86
 OY 103 VYEMRNYPFLGPGVCGYFTALFETVYCFASLITTVSVERYVALILPFRAKLOSTR 162
 DB 87 IYGSW-----VGVYVCLCITYLYOYLGINASSCSITAFTEIERYIAICHPIKQFL 141


```

FT DOMAIN 1 28 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 29 51 1 (POTENTIAL).
FT DOMAIN 52 61 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 62 83 2 (POTENTIAL).
FT DOMAIN 84 99 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 100 121 3 (POTENTIAL).
FT DOMAIN 122 144 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 145 168 4 (POTENTIAL).
FT DOMAIN 169 193 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 194 215 5 (POTENTIAL).
FT DOMAIN 216 266 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 267 288 6 (POTENTIAL).
FT DOMAIN 289 296 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 297 319 7 (POTENTIAL).
FT DOMAIN 320 412 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 3 3 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 10 10 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 376 412 YVSTKVSFDDTCLASEKNGPSSCTGYSLTAKQKXI ->
FT CONFLICT 59 59 A -> P (IN REF. 6 AND 7).
FT CONFLICT 223 223 I -> T (IN REF. 6 AND 7).
FT CONFLICT 321 321 N -> T (IN REF. 6).
FT CONFLICT 393 393 MISSING (IN REF. 3 AND 4).
SQ SEQUENCE 412 AA: 46608 MW: 52835C9849DA8A72 CRC64:

```

Query Match 18.0%; Score 393; DB 1; Length 412;
 Best Local Similarity 29.8%; Pred. No. 1,2e-17;
 Matches 111; Conservative 69; Mismatches 116; Indels 74; Gaps 13;

```

OY 18 LEDPFOKHLNSTEELYLAFLCGRSHFLPVSVVYVPIFV---GVIGNVLCVLIQH 73
DB 1 MENETVSELNQTEL-----PQOVAVALEYGVVITLVVIGLGIVGINVAVLVMT 53
OY 74 QAKKPTNYVLSLANSDDLVL--LCMPL---EYEMKRYVPLEPGVCFETALFEY 129
DB 54 KKHRRATNGVLSLAVADLVLAAGLPNTDSISYSW---VYGVCLCTITYLOYLG 108
OY 130 CPASLITTVSERYVALHPRAKLOSTRRALRILGIVGFSVLSLPNTSINGIKF 189
DB 109 INASSCSITAFTERIALICHPKAOFLCTFSAKKIIFVMAFTSYICMLNFFLDLNI 168
OY 190 HYFNGSLVPGSATC-----TVIKPMIYVNFIIQVTSFELLPMTVISLVLYMALRL 243
DB 169 SYTKDAIVI---SCGYKISRNYSPILMDFGV-----FYVPMILATVLGFIARIL 218
OY 244 -----KKKKSLEADCGNANIQRPC-----KSNYKMLFVLYVFAIC 280
DB 219 FLNPIPSDKENSKYWKNDSTHQKNMNLPTNRCFNSVSVSRKQYTKMLAVVILFALL 278
OY 281 MAPFH---IDRLFFS---FVEEMSESLAAVFNLVHVGVEFYLSAVNPITVNLSSRFQ 335
DB 279 WMPYRLLVYVNSFLSPFOENW-----FLIFCRICITVINSAINPIVYIWMGOKFR 328
OY 336 AAFQNVISFHK 347
DB 329 AAFRIKLCNCKOK 340

```

RESULT 15
 TRFR_BOVIN STANDARD; PRT; 398 AA.
 AC 046639;
 DT 20-AUG-2001 (Rel. 40, Created)
 DT 20-AUG-2001 (Rel. 40, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE THYROTROPIN-RELEASING HORMONE RECEPTOR (TRH-R) (THYROLIBERIN RECEPTOR).
 GN TRHR.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.

```

OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-HOLSTEIN;
RX MEDLINE=98151357; PubMed=9492373;
RA Takata M., Shimada Y., Ikeda A., Sekikawa K.;
RT "Molecular cloning of bovine thyrotropin-releasing hormone receptor gene."
RL J. Vet. Med. Sci. 60:123-127(1998).
CC -!- FUNCTION: RECEPTOR FOR THYROTROPIN-RELEASING HORMONE. THIS
CC RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE A
CC PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: D83964; BAA24069.1; -
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PR00751; THYROLIBRINR.
DR PROSITE: PS00237; G_PROTEIN_RECPT_FL_1; 1.
DR PROSITE: PS00262; G_PROTEIN_RECPT_FL_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
FT DOMAIN 1 28 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 29 51 1 (POTENTIAL).
FT DOMAIN 52 61 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 62 83 2 (POTENTIAL).
FT DOMAIN 84 99 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 100 121 3 (POTENTIAL).
FT DOMAIN 122 144 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 145 168 4 (POTENTIAL).
FT DOMAIN 169 193 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 194 215 5 (POTENTIAL).
FT DOMAIN 216 266 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 267 288 6 (POTENTIAL).
FT DOMAIN 289 296 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 297 319 7 (POTENTIAL).
FT DOMAIN 320 398 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 3 3 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 10 10 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 398 AA: 45165 MW: D9AF4B21A5701B8 CRC64:

```

Query Match 17.8%; Score 389.5; DB 1; Length 398;
 Best Local Similarity 29.5%; Pred. No. 1,9e-17;
 Matches 112; Conservative 73; Mismatches 112; Indels 83; Gaps 14;

```

OY 48 VSVVYVPIEV-VGVIGNVLCVLIQHQAOKTPTNYVLSLANSDDLVL--LCMPL---E 102
DB 27 VTILVLIIGLGIVGINVAVLVMTKHMRTPTNCLVSLAVADLVLAAGLPNTDS 86
OY 103 VYEMKRYVPLEPGVCFETALFEYVCFASLITTVSERYVALHPRAKLOSTRRR 162
DB 87 IYGSW---VYGVCLCTITYLOYLGINSCTSTAFTERIALICHPKAOFLCTFSR 141
OY 163 ALRILGIVGFSVLSLPNTSINGIKFHYFNGSLVPGSATC-----TVIKPMIYVNF 216
DB 142 AKKIIFVMAFTSYICMLNFFLDLNI SYTKDAIV---SCGYKISRNYSPILMDFG 197
OY 217 IQVTSFELVLPMTVISLVLYMALRL-----KKKKSLEADCGN 255
DB 198 V-----FYVPMILATVLGFIARILFLNPIPSDKENSKMKNSTHQKNLNSKTSN 251
OY 256 A--NIQRPCSKSVNKMFLVLYVFAICMAPFH---IDRLFFS---FVEEMSESLAAVFNLY 308
DB 252 RYFNSTVSVSRKQYTKMLAVVILFALLMMPYRLLVYVNSFLSPFOENW-----F 301

```


THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2002, 14:12:47 ; Search time 25.8 Seconds
(without alignments)
1225.288 Million cell updates/sec

Title: US-09-609-146-4
Perfect score: 2185
Sequence: 1 MSGMERLQNSMWTQOKLED.....ALSSEOMSRNYSFHNKKT 415

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	489.5	22.4	424	2 JH0164	neurotensin recept
2	480	22.0	378	2 T15816	hypothetical prote
3	471	21.6	418	2 S29506	neurotensin recept
4	410	18.6	416	2 S68822	neurotensin recept
5	405.5	18.6	398	2 JN0708	thyrotropin-releas
6	397	18.2	393	2 A39251	thyrotropin-releas
7	393	18.0	411	2 I56444	thyrotropin-releas
8	393	18.0	412	2 S23436	thyrotropin-releas
9	372	17.0	402	2 I56595	thyrotropin recept
10	365	16.7	369	2 B41795	neurokinin 2 recept
11	359.5	16.5	369	2 JC2083	somatostatin recept
12	358	16.4	369	2 A45291	somatostatin recept
13	357.5	16.4	399	2 S29480	bombesin recept
14	357	16.3	392	2 S65693	opioid receptor mu
15	357	16.3	398	2 A57510	opioid receptor mu
16	356.5	16.3	352	2 JE0296	thyrotropin releas
17	355	16.2	400	2 I56553	mu opiate receptor
18	352.5	16.1	380	2 S36143	kappa opioid recept
19	351.5	16.1	385	2 S55524	neurokinin 3 recept
20	351	16.1	384	2 A47249	brain-specific som
21	350.5	16.0	380	2 JC2338	kappa opioid recept
22	350.5	16.0	380	2 A55259	kappa opioid recept
23	350	16.0	384	2 JC4629	somatostatin recept
24	349	16.0	369	2 D41795	somatostatin recept
25	348.5	15.9	380	2 A48227	kappa opioid recept
26	348.5	15.9	380	2 A38271	serotonin recept
27	347.5	15.9	391	2 A41795	somatostatin recept
28	347.5	15.9	391	2 C41795	somatostatin recept
29	347.5	15.9	391	2 A39297	somatostatin recept

30	347	15.9	394	2 JC7209	galanin receptor -
31	346.5	15.9	363	2 I57940	somatostatin recept
32	346	15.8	346	2 S29248	somatostatin recept
33	345.5	15.8	452	2 A34916	neurokinin 3 recept
34	344.5	15.7	398	2 I56517	mu opioid receptor
35	343.5	15.7	367	2 I56520	g protein-coupled
36	343	15.7	366	2 S20303	neurokinin 2 recept
37	342.5	15.7	399	2 A46632	bombesin-like pept
38	341.5	15.6	380	2 JC2434	kappa opioid recept
39	341	15.6	388	2 JN0605	somatostatin recept
40	340.5	15.6	367	2 I49022	kappa opioid recept
41	340.5	15.6	367	2 JC2421	opioid receptor ho
42	339.5	15.5	384	1 S00516	neurokinin 2 recept
43	339.5	15.5	465	1 J01517	neurokinin 3 recept
44	339	15.5	398	1 J01059	neurokinin 2 recept
45	338	15.5	418	2 A46226	somatostatin recept

ALIGNMENTS

RESULT 1
JH0164
neurotensin receptor - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 17-Mar-2000
C:Accession: JH0164
R:Tanaka, K.; Masu, M.; Nakanishi, S.
Neuron 4, 847-854, 1990
A:Title: Structure and functional expression of the cloned rat neurotensin receptor.
A:Reference number: JH0164; PMID:90297956
A:Accession: JH0164
A:Molecule type: mRNA
A:Residues: 1-424 <TRAN>
C:Comment: Neurotensin receptor belongs to the family of G protein-coupled receptor.
C:Superfamily: vertebrate rhodopsin
C:Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein
F:97-121/Domain: transmembrane #status predicted <TM2>
F:144-165/Domain: transmembrane #status predicted <TM3>
F:189-210/Domain: transmembrane #status predicted <TM4>
F:236-260/Domain: transmembrane #status predicted <TM5>
F:309-330/Domain: transmembrane #status predicted <TM6>
F:348-372/Domain: transmembrane #status predicted <TM7>
F:438-442/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match	22.4%	Score 489.5;	DB 2;	Length 424;
Best Local Similarity	32.3%	Pred. No. 5.1e-34;		
Matches 107;	Conservative 75;	Mismatches 114;	Indels 35;	Gaps 9;
OY	48	VSVYVVEIVGVYIGNVLCVLIQH--QAMKPTNYVLFSLAVSDLLVLLGMPLEVV	104	
DB	67	VTAIVYLFVGVGVGVSVFTFLARKKSLQSLSTVHYHGLSLDILLAMPVELY	126	
OY	105	E-MWRNVPFLPGVGC--YFKTALFETVCFASLTSTTVSVERYVNLIPRAKIQSTR	160	
DB	127	NFTVWHPMVFGDAGCGYTF--LRDACTYATLANVASLVEKYVLIICHPFAKTLMSR	183	
OY	161	RRALRLIIGVGFVSFLPNTSINGIKFYFPGNSLVPSATCTVYKPMIYFIQVY	220	
DB	184	SRKKFKFISALWLSALLAIFMLFTMGQNR--SGDGNHPGIVCTPVDIATATKVIQVN	241	
OY	221	SPLEFLPMTIVISLVLYMALRLKKDKSLDADG-----NANTQPRCK	264	
DB	242	TFSFLEFPMVLIVILNTVIANKLTVWVHQAEQGRVCTVGNGLHSTFNMTIE-GRV	300	
OY	265	SVVK---MFLVLYVVAIGMARPHIDRLFFESV--EENESLSAAVNLVHVSGVFYI	318	
DB	301	QALRHGVLAIVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVV	360	
OY	319	SSAVNPITVNLISRRFOAFQNVISFHNKQW	349	

OY 107 WRNYPFLGPGVGC---YFKTALFEYVCFASILSTITVSEVERVAAILHPFRKLOSTRRR 163
 F:101-121/Domain: transmembrane #status predicted <TM2>
 E:101-121/Domain: transmembrane #status predicted <TM3>
 F:146-168/Domain: transmembrane #status predicted <TM4>
 F:194-215/Domain: transmembrane #status predicted <TM5>
 F:267-288/Domain: transmembrane #status predicted <TM6>
 F:297-319/Domain: transmembrane #status predicted <TM7>
 Db 153 RRLSLVWVASLGLALPMVAVIMQKHEVESADEPEPASVCVLSVSRATLQVFIQVNV 212
 OY 223 LFYLPMTVTSVLSV-----YYLML-----RLKKDSLEADG----- 254
 Db 213 VSFALPDLALFALFNGITVNHMLVLSQVPSASQVSSISRLLELSEEGGLFITWRKTL 272
 OY 255 -----NANTQPCRSKVNKMLFVLVFAICWAPFHIDRLFESEVE--WSES 300
 Db 273 SLGQVQSLVRHKDAQIRSLQNSA-QVLRALVAVVYICMLPHARLRMLCYIPDDQWTNE 331
 OY 301 LAAVFLVAVVSGVFYSSAVNPITLYNLRRFOAFOAVNISSEFKQWHS 351
 Db 332 LVDFVHYFVWVNTLFYVSSAVTPILYNAVSSSFRLFLFSLGSLGCEQHS 382

RESULT 5

thyrotropin-releasing hormone receptor - human

C.Species: Homo sapiens (man)

C.Date: 30-Sep-1993 #sequence_revision 20-Aug-1994 #text_change 20-Jun-2000

C.Accession: S40682; JN0759; S50151; S50152; J38356; JN0708

R.Marte, V.; Katsen, H.E.; Wright, M.S.; Lundell, I.; Fjeldheim, A.K.; Gabrielsen, O.S. Blochm. Biophys. Res. Commun. 195, 179-185, 1993

A.Title: Molecular cloning of a functional human thyrotropin-releasing hormone receptor.

A.Reference number: S40682; MUID:93371401

A.Accession: S40682

A.Status: preliminary

A.Molecule type: mRNA

A.Residues: 1-398 <MAT>

A.Cross-references: EMBL:X75071; NID:9404157; PIDN:CAA52965.1; PID:9404158

R.Yamada, M.; Monden, T.; Satoh, T.; Satoh, N.; Murakami, M.; Iriuchijima, T.; Kakegawa, Blochm. Biophys. Res. Commun. 195, 737-745, 1993

A.Title: Pituitary adenomas of patients with acromegaly express thyrotropin-releasing h

A.Reference number: JN0759; MUID:93384596

A.Accession: JN0759

A.Molecule type: mRNA

A.Residues: 1-398 <YAM>

A.Cross-references: GB:D16845; NID:9577631; PIDN:BA04120.1; PID:9577632

R.Hinuma, S.; Hosoya, M.; Ogi, K.; Tanaka, H.; Nagai, Y.; Onda, H. Blochm. Biophys. Acta 1219, 251-255, 1994

A.Title: Molecular cloning and functional expression of a human thyrotropin-releasing h

A.Reference number: S50151; MUID:95002135

A.Accession: S50151

A.Status: preliminary

A.Molecule type: DNA

A.Residues: 1-263 <HIN>

A.Accession: S50152

A.Status: preliminary

A.Molecule type: DNA

A.Residues: 267-398 <HI2>

R.Dublie, S.M.; Taylor, P.L.; Anderson, L.; Cook, J.; Eldine, K.A. Mol. Cell. Endocrinol. 95, R11-R15, 1993

A.Title: Cloning and functional characterisation of the human TRH receptor.

A.Reference number: 138356; MUID:94063224

A.Accession: 138356

A.Status: translated from GB/EMBL/DBJ

A.Molecule type: mRNA

A.Residues: 1-398 <RES>

A.Cross-references: EMBL:X72089; NID:9440155; PIDN:CAA50979.1; PID:9440156

C.Gene: GDB:TRHR

A.Cross-references: GDB:228955; OMIM:188545

A.Map position: 8q23-q23

C.Superfamily: adenosine receptor A1

C.Keywords: G protein-coupled receptor; receptor; transmembrane protein

F:29-51/Domain: transmembrane #status predicted <TM1>

F:62-83/Domain: transmembrane #status predicted <TM2>
 F:101-121/Domain: transmembrane #status predicted <TM3>
 F:146-168/Domain: transmembrane #status predicted <TM4>
 F:194-215/Domain: transmembrane #status predicted <TM5>
 F:267-288/Domain: transmembrane #status predicted <TM6>
 F:297-319/Domain: transmembrane #status predicted <TM7>

Query Match

18.6%; Score 405.5; DB 2; Length 398;

Best Local Similarity 29.9%; Pred. No. 6.5e-27; Mismatches 130; Indels 77; Gaps 14;

Db 48 VSVVVPPIEFV-VGVIGNVLCLVILQHOAKPTPNYYLFSLAASDLVLL-LGMP- -E 102
 OY 27 VTLLVLIICGLGICVGINVAVLVMTKHKRTPTNCLVSLAVADLMLVLAAGLPITDS 86
 Db 103 VEDMRRNYPFLGPGVGCYFKTALFEYVCFASILSTITVSEVERVAAILHPFRKLOSTRRR 162
 OY 87 IYGSW-----VGVYGCCLCTIYQLQGINASSCSITAFTERIYAIICHPIKAQFLCTFSR 141
 Db 163 ALRILGIVGCFSLFSLPNTSHIGIKFHPNSLYPGSATC-----IVIKPMIYNFI 216
 OY 142 AKRIIFVAFSLVLCMLFELDLNISTYKDAIV- -SCGYKISRNYSIYILMDG 197
 Db 217 IQVTSFLYLLPMTVTSVLYLMLRL-----KKDKSLEADGNNIOR 260
 OY 198 V-----FVVPPIILATVLGFAIRLLFPLPSPDKRENSKTKNDSTHONTLNNTSN 251
 Db 261 PC-----RKSVMKLFVLVFAICWAPFH--IDRLFEFS--FVEMSESLAAVFNLY 308
 OY 252 RCENSTVSSRKQVTKMLAVVILFALLMMPYRFLVAVNSPLSPDENW-----F 301
 Db 309 HAVSGVFYLLSAVNPIIYNLSRRFOAFONVSSFFHKQWHSQHOPLPRAORNIPL- 367
 OY 302 LFCRCITLYNSAIPNIVYLNLSQKFRFAERKLCNCKK-----PTKPNVSYALNV 354
 Db 368 -----ECHFEVELTEDIGPOFPCQSSMHNHLPALSSSE 400
 OY 355 SVIKESDHSSTELDDITVDITVDSATKVSFDDCLASE 392

RESULT 6

thyrotropin-releasing hormone receptor - mouse

C.Species: Mus musculus (house mouse)

C.Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 11-Jan-2000

C.Accession: A39251

R.Straub, R.E.; Freese, G.C.; Joho, R.H.; Gershengorn, M.C. Proc. Natl. Acad. Sci. U.S.A. 87, 9514-9518, 1990

A.Title: Expression cloning of a cDNA encoding the mouse pituitary thyrotropin-rel

A.Reference number: A39251; MUID:91088548

A.Accession: A39251

A.Molecule type: mRNA

A.Residues: 1-393 <STR>

A.Cross-references: GB:M59811; GB:M37490; NID:9202153; PIDN:AAA40480.1; PID:9202154

C.Superfamily: adenosine receptor A1

C.Keywords: G protein-coupled receptor; glycoprotein; membrane protein

Query Match

18.2%; Score 397; DB 2; Length 393;

Best Local Similarity 28.6%; Pred. No. 3.4e-26; Mismatches 110; Conservative 72; Mismatches 104; Indels 98; Gaps 15;

OY 18 LEDPFGKHLNSTE-----EY-----LAFLCGPRRSHEFLPVSVVYPIFVGY 61
 Db 1 MENDTVSEKMQTELPOAAVALEXYQVVTLLVLIICG-----LGIV 41
 OY 62 GNVLCVLILOHOAKPTPNYYLFSLAVSOLVLL-LGMP- -EYEMMRNYPFLGPGV 117
 Db 42 GNIMVAVVAKTKHMTPTNCLVSLAVADLMLVLAAGLPITDSITGSH-----YGVY 96
 OY 118 GCYFKTALFEYVCFASILSTITVSEVERVAAILHPFRKLOSTRRRRLILGIWGSVLF 177

F:254-286/Domain: transmembrane #status predicted <TM6>
 F:292-316/Domain: transmembrane #status predicted <TM7>
 F:9,22,29,32/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:115-193/Disulfide bonds: #status predicted
 F:244,343/Binding site: phosphate (Ser) (covalent) (by calmodulin-dependent kinase II) #
 F:250/Binding site: phosphate (Ser) (covalent) (by cAMP-dependent kinase) #status predicted
 F:328/Binding site: palmitate (Cys) (covalent) #status predicted

Query Match 16.4%; Score 359.5; DB 2; Length 369;
 Best Local Similarity 29.3%; Pred. No. 4,9e-23;
 Matches 103; Conservative 67; Mismatches 138; Indels 43; Gaps 12;

QY 6 KIQNASWYQKLEDPQ-----KHLNSTEYLAFLGCPRRSHFLPVSVYVPIFV 57
 DB 6 ELNGS---QPMLSPPDLNGSVATANSSNOTEPYIDLTSN-----AVLTFFYFVCI 55
 QY 58 VGVIGNVLCVLTLOHQAMKPTNYLFLSLAVSDLLVLLGMP-LEVEEMRNYPFLGCP 116
 DB 56 IGLGNTLVYVILRYAKMKTITNYILNLAIDEL-FMLGLPFLAMQVALVHMP-FGK 112
 QY 117 VGGYFTALFEYVCFASLITTVSVRYVAIILHPRAKQSTRRALRLIGIYMGSVL 176
 DB 113 AICRVVMTVDGINQFTSIFCLTVMSIDRYLAVVPIKSAKWRPRRAKMINVAVGSSL 172
 QY 177 FSLPNTSIHGKHFYPNGSLVPGSATCTVIRP---MVIYNFLIQVTSFLYLLPMTVI 232
 DB 173 VILPIMIVAGLRNQM-----GRSCTINMPGEGSAWGTGFI-YAFLGLFLVPLTII 224
 QY 233 SVLYLMALRLKDKSLDEAGNANIQPCRSVNMKLFVLVLAFCAPPHIDL-FF 291
 DB 225 CLCYELRIIVKVS---SGIRVGSKRKKSEKVKTRMVSIVAVFIQCWLPFIYFNVSV 280
 QY 292 SFVEEMSESLAAVENLVHVGVFYLSAVNPIYVNLSSRRQAQFQNI 342
 DB 281 SVAISPTPALKGMFDEVYVLT-----YANSCANPLIYALFSLDNKFSQNVL 327

RESULT 12
 A45291

somatostatin receptor, somatotropin release-inhibiting factor receptor, SRIF receptor -
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 29-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jul-2000
 C:Accession: A45291
 R:Kluxen, F.W.; Bruns, C.; Lubbert, H.
 Proc. Natl. Acad. Sci. U.S.A. 89, 4618-4622, 1992
 A:Title: Expression cloning of a rat brain somatostatin receptor cDNA.
 A:Reference number: A45291; MUID:92262491
 A:Accession: A45291
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-369 <KLU>
 A:Cross-references: GB:M93273; NID:q207026; PIDN:AAA42165.1; PID:q207027
 C:Note: sequence extracted from NCBI backbone (NCBIN:102315, NCBIPI:102316)
 C:Superfamily: Vertebrate rhodopsin
 C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 16.4%; Score 358; DB 2; Length 369;
 Best Local Similarity 30.1%; Pred. No. 6,5e-23;
 Matches 97; Conservative 65; Mismatches 128; Indels 32; Gaps 10;

QY 27 NSTEYLAFLGCPRRSHFLPVSVVYVPIFVGVIGNVLCVLTLOHQAMKPTNYLFLS 86
 DB 32 NQTEPYDMTSN-----AVLTFFYFVVCVGLCGNTLVYILIRAKMKTITNIIYLN 84
 QY 87 LAVSLLVLLGMP-LEVEEMRNYPFLGCPVCCYKTALEFVCFASLITTVSVRY 145
 DB 85 LAIADEL-FMLGLPFLAMQVALVHMP-FGKAICRVVMTVDGINQTSIFCLTVMSIDRY 141
 QY 146 VALHPRAKLQSTRRALRLIGIYMGSVLFLPNTSIHGKHFPRNSLVPGSATCT 205
 DB 142 LAVVPIKSAKWRPRRAKMINVAVGVSLLVLPIMITYAGLRNQM-----GRSCT 194

QY 206 VIKP---MVIYNFLIQVTSFLYLLPMTVISVLYLMALRLKDKSLDEAGNANIQP 261
 DB 195 INMPGEGSAWGTGFI-YAFLGLFLVPLTIIICLYLFIITIKVS---SGIRVGSKRKK 249
 QY 262 CRKSVNMKLFVLVLAFCAPPHIDL-FFSFVEEMSESLAAVENLVHVGVFYLS 320
 DB 250 SEKKVTRMVSIVAVFIFCWLFPFIYFNVSSVAISPTPALKGMFDEVYVLT-----YANS 305
 QY 321 AVNPIYVNLSSRRQAQFQNI 342
 DB 306 CANPILYALFSLDNKFSQNVL 327

RESULT 13

S29480
 bombesin receptor - guinea pig
 C:Species: Cavia porcellus (guinea pig)
 C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 16-Jul-1999
 C:Accession: S29480
 R:Gumbler, V.; Akhundova, A.; Buechner, H.; Fahrenholz, F.
 submitted to the EMBL Data Library, July 1992
 A:Description: Molecular cloning of a new bombesin receptor subtype.
 A:Reference number: S29480
 A:Accession: S29480
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-399 <GOR>
 A:Cross-references: EMBL:X67126; NID:g49545; PIDN:CAA47605.1; PID:g49546
 C:Superfamily: endothelin receptor B
 C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 16.4%; Score 357.5; DB 2; Length 399;
 Best Local Similarity 28.5%; Pred. No. 7,8e-23;
 Matches 85; Conservative 68; Mismatches 134; Indels 11; Gaps 4;

QY 48 VSVVYVPIFVGVIGNVLCVLTLOHQAMKPTNYLFLSLAVSDLLVLLGMPLE-----V 103
 DB 50 IYITVAVIISVGLNAILKVEFKTSMQTVNRIFITSLALGDLILLTCVVDATHYL 109
 QY 104 YEMWRNYPFLGPGCYEKTALFEYVCFASLITTVSVRYVAIILHPRAKLQSTRRA 163
 DB 110 AEGW----LFGRIQCKVLSFIRLTSGVSVFTLTISADRYAKAVVPLERQPSNALKT 164
 QY 164 LRLIGIYMGSVLFLPNTSIHGKHFYPNGSLVPGSATCTVIRKPMIYNFLIQVTSFL 223
 DB 165 CANAGCIVMISMIFALPEALFSNVHTLRDPNKNMTSEMAFYVSEKLOEIHALLSEFLV 224
 QY 224 FYLLPMTVISVLYLMALRL-KDKDKSLADE-NNANIQPCRSVNMKLFVLVLAFCW 281
 DB 225 FYIIPLSIISVYSLIARLTLYKSTLNIPTEQSHARQVSRKRRIKTYLVLAALFALCW 284
 QY 282 APPIHIDLFFSFVEEMSESLAAVENLVHVGVFYLSAVNPIYVNLSSRRQAQFQ 339
 DB 285 LPMHLLNLVHSHFKAYEDSSAIFHTVITPSRYLAFNSCVNPFALYWLSTKQKQFK 342

RESULT 14

S65693
 opioid receptor mu variant MOR1A - human
 C:Species: Homo sapiens (man)
 C:Date: 12-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 05-Nov-1999
 C:Accession: S65693; S51216
 R:Barre, L.A.; Mansson, E.; Yang, D.
 submitted to the EMBL Data Library, July 1994
 A:Description: Expression of two variants of the human mu opioid receptor mRNA in SK-
 A:Reference number: S65693
 A:Accession: S65693
 A:Molecule type: mRNA
 A:Residues: 1-392 <BAR>
 A:Cross-references: EMBL:U12569; NID:g607911; PIDN:AAB60354.1; PID:g607912
 R:Barre, L.A.; Mansson, E.; Yang, D.

FEBS Lett. 354, 213-216, 1994
 A:Title: Expression of two variants of the human mu oploid receptor mRNA in SK-N-SH cell
 A:Reference number: S51215; MUID:95046336
 A:Accession: S51216
 A:Molecule type: mRNA
 A:Residues: 387-392 <BAW>
 C:Superfamily: vertebrate rhodopsin

Query Match 16.3%; Score 357; DB 2; Length 392;
 Best Local Similarity 27.8%; Pred. No. 8.5e-23;
 Matches 100; Conservative 65; Mismatches 125; Indels 70; Gaps 12;

QY 11 SWIQOKLEDPKHLNSTEYLAFLGPRSRHF-----FLPVSVVY 53
 DB 30 SWV-----NLSHLDGSLDPCGPNRDLGGDSLCPTGSPSMATITIALIS 78
 QY 54 PLFVGVGNVLVCLVLIHQAMKPTNYVLFSLAVSDLLVLLGMPLE-VYEMMRNYPE 112
 DB 79 IVCVGLFENFLVMYIVRYTKMTATNIFYENLALDALATST-LPQSVNYLMGTWP- 136
 QY 113 LGPVCCKRTALFEVVCASISITTVSEVERVALHPRAKLOSTRRALILGIWVG 172
 DB 137 -FGTILCKIVISIDYNNMFTSLFTLCTMSVDRIYAVCHPKALDFRPRNAKLIINCMNI 195
 QY 173 FSVLFLPMTSIGHKEFHFPNGSLVPGSATCTVI--KPMIYNFTIOVTSFLF-YLLPM 229
 DB 196 LSAIGLPPWFATITTKYR-----QGSIDCTLTFESHPTWYENLKIICVFIFAFIMPY 247
 QY 230 TWISVLVYLMALRLKKDKSL-ADENANIQRCRKSVMKMLFVLVFAICMAPPHIDR 288
 DB 248 LITVVCYGLMILRLKSVRLSGSKERDLNR-----ITRMVLVVAVFVCTPIHIYV 302
 QY 289 LPESFPE-----EMESLAAVPNLVHVVGVEFYSSANPIIYNLSRRFOAFQ 339
 DB 303 IKAALVTIDETTFQTVSMHFCLIA-----LGYNSCLNPVLYAFLEDFENFRCFR 350

RESULT 15

A57510
 mu oploid receptor - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 08-Dec-1995 #sequence_revision 08-Dec-1995 #text_change 24-Nov-1999
 C:Accession: A57510; I48665; S66513; I49300
 J.Kaufman, D.L.; Keith Jr., D.E.; Anton, B.; Tian, J.; Magendzo, K.; Newman, D.; Tran, T.
 J. Biol. Chem. 270, 15877-15883, 1995
 A:Title: Characterization of the murine mu oploid receptor gene.
 A:Reference number: A57510; MUID:95318184
 A:Accession: A57510
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-398 <KAU>
 A:Cross-References: GB:U19380
 R.Mlin, B.H.; Augustin, L.B.; Felsheim, R.F.; Fuchs, J.A.; Loh, H.H.
 Proc. Natl. Acad. Sci. U.S.A. 91, 9081-9085, 1994
 A:Title: Genomic structure analysis of promoter sequence of a mouse mu oploid receptor
 A:Reference number: I48665; MUID:94377496
 A:Accession: I48665
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-398 <RES>
 A:Cross-References: EMBL:U10561; NID:9555696; PIDN:AAB60673.1; PID:g565069
 R.Rosel, G.C.; Pan, Y.X.; Brown, G.P.; Pasternak, G.W.
 FEBS Lett. 369, 192-196, 1995
 A:Title: Antisense mapping the MOR-1 oploid receptor: evidence for alternative splicing
 A:Reference number: I49300; MUID:95377399
 A:Accession: S66513
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: mRNA
 A:Residues: 1-398 <ROS>
 A:Cross-References: EMBL:U26915; NID:g1055230; PIDN:AAA81170.1; PID:g1055231
 A:Note: The nucleotide sequence was submitted to the EMBL Data Library, November 1995
 C:Genetics:

A:Gene: MOR-1
 A:Introns: 95/2; 213/1; 386/3
 C:Superfamily: vertebrate rhodopsin
 C:Keywords: alternative splicing; G protein-coupled receptor; glycoprotein; transmem

Query Match 16.3%; Score 357; DB 2; Length 398;
 Best Local Similarity 27.5%; Pred. No. 8.6e-23;
 Matches 105; Conservative 67; Mismatches 144; Indels 66; Gaps 14;

QY 37 CGPRR-----SHEFLP-----VSVVYPIFVGVGNVLVCLVLIHQAMKTP 79
 DB 43 CGPNRTGLGSHSLCPTGSPSMATITIALISIVCVGLFENFLVMYIVRYTKMTA 102
 QY 80 TNYVLSLAVSDLLVLLGMPLE-VYEMMRNYPELGPGCYKRTALFEVVCASISIT 138
 DB 103 TNYIYFNLALDALATST-LPQSVNYLMGTWP--FGNITCKIVISIDYNNMFTSLFTLC 159
 QY 139 TVSEVERVALHPRAKLOSTRRALILGIWVGFSVLFLPMTSIGHKEFHFPNGSLV 198
 DB 160 TMSVDRIYAVCHPKALDFRPRNAKIYVNCNIISSAIGLPWFMTTKYR----- 211
 QY 199 PGSATCTVI--KPMIYNFTIOVTSFLF-YLLPMVIVSVLYLMALRLKKDKSL-ADRG 254
 DB 212 QGSIDCTLTFESHPTWYENLKIICVFIFAFIMPYLITVVCYGLMILRLKSVRLSGSKER 271
 QY 255 NANIQRCKSVKMLFVLVFAICMAPPHIDRLEFSFPE-----EMESLAAYF 305
 DB 272 DRNLRR-----ITRMVLVVAVFVCTPIHIYVLIKALITIDETTFQTVSMHFCLIA--- 323
 QY 306 NLVHVVGVEFYSSANPIIYNLSRRFOAFQNY---ISSFRKOWHS-----QHPOLP 358
 DB 324 -----LGYNSCLNPVLYAFLEDFENFRCFRCECIPITSSITIEQNSAIRQNRHNP 374
 QY 359 PAQRNIFLTCHEFVELTEDIGP 380
 DB 375 STANTVDRTNHOLENLEAETAP 396

Search completed: April 22, 2002, 14:17:43
 Job time: 296 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: Apr11 22, 2002, 15:47:04 ; Search time 116.57 Seconds
(without alignments)
2521.819 Million cell updates/sec

Title: US-09-609-146-3

Perfect score: 1298

Sequence: 1 agggagagcctcagcctctg.....tttcagagcctcctctc

Scoring table: IDENTITY_NUC

Searched: Gapop 10.0 , Gapept 1.0

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/lna/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/lna/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/lna/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/lna/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/lna/PTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/lna/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	149.6	11.5	1088	4	US-09-077-675A-6
2	144.6	11.1	1063	4	US-09-077-675A-1
3	134.8	10.4	1092	4	US-09-077-675A-15
4	120.4	9.3	1122	4	US-09-077-675A-9
5	117.6	9.1	836	4	US-09-077-675A-11
6	112.4	8.7	1029	4	US-09-077-675A-4
7	112.4	8.7	1161	1	US-08-086-439C-2
8	112.4	8.7	1161	1	US-08-434-877-2
9	112.4	8.7	1367	3	US-08-475-742-3
10	112.4	8.7	1370	1	US-08-056-051-1
11	112.4	8.7	1370	1	US-07-928-611-17
12	112.4	8.7	1370	2	US-08-487-811A-17
13	112.4	8.7	1370	4	US-09-060-694-11
14	112.4	8.7	1370	5	PCT-US93-07370-17
15	112.4	8.7	1466	1	US-08-056-051-3
16	112.4	8.7	1466	1	US-07-928-611-19
17	112.4	8.7	1466	2	US-08-487-811A-19
18	112.4	8.7	1466	4	US-09-060-694-19
19	112.4	8.7	1466	5	PCT-US93-07370-19
20	112.4	8.7	1610	1	US-08-056-051-5
21	112.4	8.7	1610	1	US-07-928-611-21
22	112.4	8.7	1610	2	US-08-487-811A-21
23	112.4	8.7	1610	4	US-09-060-694-21
24	112.4	8.7	1610	5	PCT-US93-07370-21
25	110	8.5	1529	3	US-08-838-876A-3
26	110	8.5	1529	4	US-09-472-880-3
27	109.2	8.4	3129	4	US-09-077-675A-14

28	97.6	7.5	1575	3	US-08-858-876A-1	Sequence 1, Appl
29	97.6	7.5	1575	4	US-09-472-880-1	Sequence 1, Appl
30	97.2	7.5	283	4	US-08-993-088A-4	Sequence 4, Appl
31	93.2	7.2	1164	4	US-08-993-088A-6	Sequence 6, Appl
32	88.8	6.8	1601	1	US-08-722-001-7	Sequence 7, Appl
33	88.8	6.8	1997	1	US-08-722-001-27	Sequence 27, Appl
34	88.8	6.8	2004	1	US-08-722-001-11	Sequence 11, Appl
35	88.6	6.8	1639	1	US-08-334-698-5	Sequence 5, Appl
36	88.6	6.8	1639	1	US-08-228-932-5	Sequence 5, Appl
37	88.6	6.8	1639	1	US-08-468-939-5	Sequence 5, Appl
38	88.6	6.8	1639	2	US-08-406-855A-5	Sequence 5, Appl
39	88.6	6.8	1639	2	US-08-722-190-5	Sequence 5, Appl
40	88.6	6.8	1639	3	US-08-244-354-5	Sequence 5, Appl
41	88.6	6.8	1639	3	US-09-206-899-5	Sequence 5, Appl
42	88.6	6.8	1639	5	PCT-US95-04203-5	Sequence 5, Appl
43	87.6	6.7	1342	3	US-08-832-389-1	Sequence 1, Appl
44	87.6	6.7	1342	4	US-09-372-498-1	Sequence 1, Appl
45	87	6.7	1205	1	US-08-417-103-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1
US-09-077-675A-6
; Sequence 6, Application US/09077675A
; Patent No. 6242199
GENERAL INFORMATION:
; APPLICANT: Pal, Lee-Yuh
; APPLICANT: Feigner, Scott C.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Peng, Sheng-Shung
; APPLICANT: Van Der Ploeg, Leonardus H.T.
; TITLE OF INVENTION: RECEPTOR ASSAY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077,675A
; FILING DATE: 3-JUN-1998
CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19590P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
TELEX:
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1088 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-09-077-675A-6
Query Match 11.5%; Score 149.6; DB 4; Length 1088;


```

;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077.675A
; FILING DATE: 3-JUN-1998
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19590P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
; TELEX:
;
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1029 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
;
; US-09-077-675A-4

```

```

Query Match      8.7%  Score 112.4; DB 4; Length 1029;
Best Local Similarity 52.1%  Pred. No. 1.1e-18;
Matches 301; Conservative 0; Mismatches 271; Indels 6; Gaps 2;

```

```

Oy 303 ctggtctcctccttggaatgccccctggaagtctatgagatgtggcgcaactacccttc 362
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3 ctgctcatcttctctctgcatgacctgacctgcttgcctctggcagataccggcccttg 62
Oy 363 ttgtctcggccgttggtctgtacttcaagaagccctcttgagaacgtgtgtctgcgc 422
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 63 AACCTTCGGCGACCTCCTCTGCAAACTCTTCCAAATTCGTAGTAGAGCTGCACATACGCC 122
Oy 423 tcaatcctcagaatacaacacgttcagcgttgagcgttaacgttgccatccctaacccgttc 482
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 123 ACGGTGCTTCACATCAGACACCGGTGAGCCGTGAGCCGTACTTCGCCATCTGCTTCCACATC 182
Oy 483 cgcgcgaacacacagagacacccgcgcgcgcgcctcgaagatccctcgtgagatctgtcgggc 542
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 183 CGGGCCCAAGGTGTGTGTGTCACCAAGGGGGGGGTGAAGCTGTATCTTCTGTCATCTGGGCC 242
Oy 543 ttctcgtgtctctctcctcctcccaacacagatccatcgtacatcaagttcactacttc 602
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 243 GTGGCCTTCTGAGCGCCGGCCCATCTTGTGTAGTGGGGTGGAGACGAGAACGGCC 302
Oy 603 cccaatgggtcccttggtcccaagttcgcacactgtaagttcaagcccatgtgagac 662
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 303 ACCGACCTTGGGACACCAACAGATGCCGCCACCGAGTTTGGGGTCTGTGAGCTG 362
Oy 663 taacaattcatatcccaagttcaacctcctcctatctactcctccctcccaatgactgtatc 722
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 363 CTCACGGCTATGGTGTGGGT---GTCCAGCATCTTCTTCTCTCTCTCTCTCTCTCTCTCTC 419
Oy 723 agtgtctctactactcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 782
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 420 ACGGTCTCTCAGTCTCATCTCGCAAGAGCTGTGGCGAGAGCGCGCGGCTGTCTGTCTGTC 479
Oy 783 gaagggaatgcaaatctcaaaagaccctcagagaatactcaagaagtgtcgttctgc 842
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 480 GTGGGCTCTCCTCGTCAGGACCAAGAACACCAAGCAACCGGTGAAA---ATGCTGCTGTGA 536
Oy 843 ttgtctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 902
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 537 GTGGTGTGCTTCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTC 596
Oy 903 ttgttggaagatggaatggaatccctcgtgctgtgtgtcaaac---tcgtccaatgtgtgtg 959
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 597 AAATCTTGTAGCGCTGCTCTGTGAGATTGCTCAGATCAGCCAGTACGAACTGCTGAG 656
Oy 960 tcaagtgctctctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 1019
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 657 TCCCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTC 716
Oy 1020 cgcgcctcctcagcagcat 1039
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 717 AGAAGTACCGGGTGGCAGT 736

```

```

RESULT 6
US-09-077-675A-4
; Sequence 4, Application US/09077675A
; Patent No. 6242199

```

GENERAL INFORMATION:

```

; APPLICANT: Pal, Lee-Yuh
; APPLICANT: Feighner, Scott C.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Pong, Sheng-Shung
; APPLICANT: Van Der Ploeg, Leonardus H.T.
; TITLE OF INVENTION: RECEPTOR ASSAY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900

```

```

Oy 467 catctacacccgtctcgcgcaaacctcagagacacccgcgcgcgcgcctcagaatcct 526
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 591 CATCTGCTTCCGCGCTGCGGCAAGGTAGTGACACCAAGGCGCGGTAAAGCTGGTAT 650
Oy 527 cgcgcatcgtctggggtctctcgtgtctctcctcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 586
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 651 CTTGTGATCTGGGCGGTGGCTTGTGACAGCGCGGCGCATCTTGTGCTGTGTGGTGGT 710
Oy 587 caagttcactacttccccaatgggtccctcgtgtcccaagttcgcgcgcgcgcgcgcgcgcgcgc 646
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 711 GAGGATGATAGGAGCAGTACCTTGGGACACCAACAGAGTGGCGGACAGAGTTCGC 770
Oy 647 caagcccatgtgatatcaaatltcaatcaaggtgaacccctcctcctcctcctcctcctcctcct 706
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 771 CGTGGCTCCGGGCTGCTTACGCTATGCTGTGGGTGCTCAGAGTCTTCTCT---TCT 827
Oy 707 cccaatgactgtatcagtgtcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 744
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 828 GCCTGTCTTCTGCTCATCTGTGCTCTATAGCCATCG 865

```


Db	292	GGCGCGCTGGCTGTGAGGCGCCCGCGCTGTGGACGCCCTCAATGGCCATGAGACGTTCATCTG	351
Oy	414	tgcttcgcctccatccctcagaacacacacgcgtacgctgtagcgctacgtgycacata	473
Db	352	TGCACCGGCTCTCATCTTCAACCTGTGGCCATCAAGCTGACAGATGTCGTGGCCGTGGCC	411
Oy	474	caaccggttcgcgcgcaaaactgcagaagacacccgcgcgcgcgcctccagatccctgcgc	533
Db	412	GTGCGCGCTGCCTCAACACCGGACAGGTGGAGCGCGCGGACGTGCTGTATCGGCGCC	471
Oy	534	gtctgggcgtctctccgtctctctccctggcacaacccagatcacatgtgcatcaagt	593
Db	472	ACGTGCGCTGCTGTCCGCGCGCGCTGGCGCGCCCGCTACTGTGCGGCTCAACAGAGCTGCGC	531
Oy	594	caactactccccaatggatccc	615
Db	532	GGCCCGACCCCGCCGCTGTGGC	553

RESULT

US-08-475-742-3

```

: Sequence 3, Application US/08475742
: Patent No. 6121015
: GENERAL INFORMATION:
: APPLICANT: O'Malley, Karen L
: APPLICANT: Todd, Richard D
: TITLE OF INVENTION: Gene Encoding the Rat Dopamine D4 Receptor
: FILE REFERENCE: WU 102 CON DIV
: CURRENT APPLICATION NUMBER: US/08/475,742
: EARLIER FILING DATE: 1995-06-07
: EARLIER APPLICATION NUMBER: US 08/261,293
: EARLIER FILING DATE: 1994-06-16
: EARLIER APPLICATION NUMBER: US 08/014,013
: EARLIER FILING DATE: 1993-01-28
: NUMBER OF SEQ ID NOS: 16
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 3
: LENGTH: 1367
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc.feature
: LOCATION: (1)..(1367)
: OTHER INFORMATION: D4 Dopamine Receptor cDNA
: PUBLICATION INFORMATION:
: AUTHORS: Bunzow, J. R.
: AUTHORS: Van Tol, H. H.
: TITLE: Cloning of the gene for a human dopamine D4 receptor
: TITLE: with high affinity for the antipsychotic clozapine
: JOURNAL: Nature
: VOLUME: 350
: PAGES: 610-614
: DATE: 1991
: US-08-475-742-3

```

Query Match	8.7%	Score 112.4	DB 3	Length 1367
Best Local Similarity	53.4%	Pred. No. 1.3e-18		
Matches 236: Conservative	0	Mismatches 206	Indels 0	Gaps 0

[illegible][illegible]

RESULT 10

US-08-056-051-1

```

1 Sequence 1 Application US/08056051
2 Patent No. 5516683
3 GENERAL INFORMATION:
4 APPLICANT: Grandy, David K
5 APPLICANT: Bunzow, James R
6 APPLICANT: Clevell, Olivier
7 APPLICANT: Van Tol, Hubert H.-M.
8 TITLE OF INVENTION: A No. 5516683el Human Dopamine Receptor and Uses
9 NUMBER OF SEQUENCES: 6
10 CORRESPONDENCE ADDRESS:
11 ADDRESSEE: Allegretti & Witcoff, Ltd.
12 STREET: 10 South Wacker Drive, Suite 3000
13 CITY: Chicago
14 STATE: Illinois
15 COUNTRY: USA
16 ZIP: 60606
17 COMPUTER READABLE FORM:
18 MEDIUM TYPE: Floppy disk
19 COMPUTER: IBM PC compatible
20 OPERATING SYSTEM: PC-DOS/MS-DOS
21 SOFTWARE: PatentIn Release #1.0, Version #1.25
22 CURRENT APPLICATION DATA:
23 APPLICATION NUMBER: US/08/056,051
24 FILING DATE: 19930429
25 CLASSIFICATION: 435
26 ATTORNEY/AGENT INFORMATION:
27 NAME: No. 5516683nan, Kevin E
28 REGISTRATION NUMBER: 35,303
29 REFERENCE/DOCKET NUMBER: 90,1092-C
30 TELECOMMUNICATION INFORMATION:
31 TELEPHONE: 312-715-1234
32 TELEFAX: 312-715-5317
33 TELEX: 910-221-5317
34 INFORMATION FOR SEQ ID NO: 1:
35 SEQUENCE CHARACTERISTICS:
36 LENGTH: 1370 base pairs
37 TYPE: NUCLEIC ACID
38 STRANDEDNESS: single
39 TOPOLOGY: linear
40 MOLECULE TYPE: cdna
41 FEATURE:
42 NAME/KEY: 5'UTR
43 LOCATION: 1..103
44 FEATURE:
45 NAME/KEY: 3'UTR
46 LOCATION: 1268..1370
47 FEATURE:
48 NAME/KEY: CDS
49 LOCATION: 104..1267
50 US-08-056-051-1

```


OY	474	caccggttcacggcgcgcaaacatgacagacacccgggaccgggccccttcagatccttcgycatc	553
Db	518	gtggccggcgccctcTCAACACGGGCAAGGAGGGAGCGCGGAGGTGTGTCTATCTGGGGCC	577
OY	534	gtctggggctctctcgtgctctctcctcctgcacaaacagacatcgaatggatcatcaagtctc	593
Db	578	ACGTTGGGCTGTGTCCGGGGGGCGTGGGGAGGCCCTACTGTGGCGCTCAACAGCAGTGGCC	637
OY	594	cactacttcgccaaatgggtccc	615
Db	638	GGCCGCGACCCCGCCGTGTGCC	659

RESULT 14
PCT-US93-07370-17

```

1  APPLICANT:
2  TITLE OF INVENTION:  A Novel Human Dopamine Receptor and Uses
3  NUMBER OF SEQUENCES:  22
4  COMPUTER READABLE FORM:
5  MEDIUM TYPE:  Floppy disk
6  COMPUTER:  IBM PC compatible
7  OPERATING SYSTEM:  PC-DOS/MS-DOS
8  SOFTWARE:  PatentIn Release #1.0, Version #1.25 (EPO)
9  CURRENT APPLICATION DATA:
10 APPLICATION NUMBER:  PCT/US93/07370
11 INFORMATION FOR SEQ. ID NO:  17:

```

Query Match	8.7%	Score	112.4	DB	5	Length	1370
Best Local Similarity	53.4%	Pred. No.	1.3e-18				
Matches	236	Conservative	0	Mismatches	206	Indels	0
						Gaps	0

Accession	Sequence	Position
OY	174 gtcgggtatgcgaacattttctgtgtggtgggtcattctgcaaatgcctgggtgtctccctgtg	2333
Db	218 GTGGGGGGGCTGCTCATGTGGGGCGGTCTCTCCGGGGAACTCGCTGCTGCTGAC	2777
OY	234 attctgcagcacaagcgtatgaagacgcccacaactactactctctcaagccttgcgtc	2933
Db	278 GTGGCCACCGAGCGCGCCCTGCACAGCCGCCCAACTCTTTCATCGTAGACCTGGCGGC	3373
OY	294 tctgaacctccctgggtcctctctcttggaatggcccttggaagttctatctgagatgtgctgaac	3533
Db	338 GCCGACCTCTCTCTGCTCTCTGTGTGCTGGCTCTTTCGTTCACTCCGAGTCCAGGTT	3973
OY	354 taaccttctctgttcggcgccgtggcgctgaactcaagaagcgccctctatgagacgt	4133
Db	398 GCGCGGTGGCTGCTGAGCGCCCGCTGTGGCAGCCCTCATAGCCATGAGATCATCTG	4573
OY	414 tgcctgcctccatcctctcaagcatcaacacggtcagcggtgagacgctacgtgcacatcta	4733
Db	458 TGCACCGCTTCATCTTCAACTGTGGGCCATCAGCGGAGCAAGTTCTGGGCCGTGGCC	5173
OY	474 caaccgcttcgagccaaactgcagagaccccgcgccggcgccctcagagatcttcggcatc	5333

Db 518 GTGCGCGGCCTCAACACGGCAGGGGAGGCGCGGCGAGCTGCTCATATGAGCGCC 577

QY 534 gtctggagatctctcgtgcgtcttcctcgtcccaacaacagatcatcatgaattc 593

Db 578 ACGTGGCGTGTCTCCGCGCGCGTGGCGGCCCTACTGTGGGCTCAACAGAGTGGCG 637

QY 594 cactactctcccaatgagtgctcc 615

Db 638 GGCGCGACCGCGCGGTGTGGC 659

RESULT 15 -
US-08-056-051-3
; Sequence 3, Application US/08056051

```

1  GENERAL INFORMATION:
2  APPLICANT: Grandy, David K
3  APPLICANT: Bunzow, James R
4  APPLICANT: Civeill, Olivier
5  APPLICANT: Van Tol, Hubert H.-M.
6  TITLE OF INVENTION: A No. 5516683e1 Human Dopamine Receptor and Uses
7  NUMBER OF SEQUENCES: 6
8  CORRESPONDENCE ADDRESS:
9  ADDRESSEE: Allegretti & Witcoff, Ltd.
10 STREET: 10 South Wacker Drive, Suite 3000
11 CITY: Chicago
12 STATE: Illinois
13 COUNTRY: USA
14 ZIP: 60606
15

```

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: No. 5516683nan, Kevin E
 REGISTRATION NUMBER: 35,303
 REFERENCE/DOCKET NUMBER: 90, 1092-C-
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 312-715-1000
 TELEFAX: 312-715-1234
 TELECY: 910-221-5317
 INFORMATION FOR SER ID NO: 3:

Query Match	8.7%	Score 112.4	DB 1	Length 1466
Best Local Similarly	53.4%	Pred. No. 1.3e-18		
Matches 236	Conservative	0	Mismatches 206	Indels 0
				Gaps 0

Oy 174 ggggtgtaatgcgaattttctgctggggcgaattgcaatgcccgtgttccttgatg 233

Db 218 gtggggggcgctgcctcatccggcgctgccggggaactcgctgctggcgtgagc 277

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 22, 2002, 14:13:46 ; Search time 2258.39 Seconds
(without alignments)
6176.090 Million cell updates/sec

Title: US-09-609-146-3

Perfect score: 1298
Sequence: 1 agggagagctcagccttgg.....ttcagagctgactctctc 1298

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 11351937 seqs, 537289281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estom:*
5: em_estcpl:*
6: em_estlba:*
7: em_estlro:*
8: em_estlov:*
9: em_hic:*
10: gb_estl:*
11: gb_estl2:*
12: gb_hic:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rtd:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	378	29.1	555	10	BE751626 203809 MA
2	339	26.1	929	11	BE871983 601448032
3	302	23.3	414	11	MA5474 yv59b04.r1
4	299.2	23.1	465	10	BE751310 203132 MA
5	211.2	16.3	941	13	CNS036WH
6	123.4	9.5	450	11	RL3890
7	116.4	9.0	959	13	CNS04GXP
8	113	8.7	379	11	HL3359
9	107.8	8.3	817	13	CNS01XJ2
10	107	8.2	500	11	BE603623
11	95.4	7.3	448	11	BF323227
12	95.2	7.3	975	13	CNS039RV

RESULT	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE	COMMENT
1	BE751626	203809 MARC 280V Bos taurus CDNA 5', mRNA sequence.	BE751626	1	GI:10165618	EST.	COW.	1	Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T., Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett,G.L., Heaton,M.P., Laegreid,W.M., Rohrer,G.A., Chitko-McKown,C.G., Pettea,C., Holt,I., Karanycheva,S., Liang,F., Quackenbush,J. and Keefe,J.W.	Sequence evaluation of four pooled-tissue normalized bovine CDNA libraries and construction of a gene index for cattle	Genome Res. 11 (4), 626-630 (2001)	21180013	Contact: Smith TPL USDA, ARS, US Meat Animal Research Center PO Box 166, Clay Center, NE 68933-0166, USA Tel: 402 762 4366 Fax: 402 762 4390 Email: smith@mail.marc.usda.gov Single pass sequencing. Bases called and alt-trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -mismatch 12 options. PCR Primers FORWARD: AGGAACAGCTATGACCAT BACKWARD: GTTTCACGTCACGACG Plate: 44 row: 0 column: 6
2	BE751626	203809 MARC 280V Bos taurus CDNA 5', mRNA sequence.	BE751626	1	GI:10165618	EST.	COW.	1	Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T., Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett,G.L., Heaton,M.P., Laegreid,W.M., Rohrer,G.A., Chitko-McKown,C.G., Pettea,C., Holt,I., Karanycheva,S., Liang,F., Quackenbush,J. and Keefe,J.W.	Sequence evaluation of four pooled-tissue normalized bovine CDNA libraries and construction of a gene index for cattle	Genome Res. 11 (4), 626-630 (2001)	21180013	Contact: Smith TPL USDA, ARS, US Meat Animal Research Center PO Box 166, Clay Center, NE 68933-0166, USA Tel: 402 762 4366 Fax: 402 762 4390 Email: smith@mail.marc.usda.gov Single pass sequencing. Bases called and alt-trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -mismatch 12 options. PCR Primers FORWARD: AGGAACAGCTATGACCAT BACKWARD: GTTTCACGTCACGACG Plate: 44 row: 0 column: 6

ALIGNMENTS

ALIGNMENT	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE	COMMENT
1	BE751626	203809 MARC 280V Bos taurus CDNA 5', mRNA sequence.	BE751626	1	GI:10165618	EST.	COW.	1	Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T., Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett,G.L., Heaton,M.P., Laegreid,W.M., Rohrer,G.A., Chitko-McKown,C.G., Pettea,C., Holt,I., Karanycheva,S., Liang,F., Quackenbush,J. and Keefe,J.W.	Sequence evaluation of four pooled-tissue normalized bovine CDNA libraries and construction of a gene index for cattle	Genome Res. 11 (4), 626-630 (2001)	21180013	Contact: Smith TPL USDA, ARS, US Meat Animal Research Center PO Box 166, Clay Center, NE 68933-0166, USA Tel: 402 762 4366 Fax: 402 762 4390 Email: smith@mail.marc.usda.gov Single pass sequencing. Bases called and alt-trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -mismatch 12 options. PCR Primers FORWARD: AGGAACAGCTATGACCAT BACKWARD: GTTTCACGTCACGACG Plate: 44 row: 0 column: 6

FEATURES	Seq primer: ATTTAGGTCACACTATAC	Location/Qualifiers
SOURCE	1..535	/organism="Bos taurus" /db_xref="taxon:9913" /clone_lib="MARC 2B0V" /tissue_type="pooled" /lab_host="DH10B" /note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI; Library made from pooled tissue from testis, thymus, semitendinosus muscle, longissimus muscle, pancreas, adrenal, and endometrium."
BASE COUNT	100 a 165 c 134 g 156 t	
ORIGIN		
Query Match	29.1%	Score 378; DB 10; Length 555;
Best Local Similarity	86.5%	Pred. No. 1.6e-81;
Matches 417; Conservative	0; Mismatches 65; Indels	0; Gaps 0;
OY	1 agggagagctcagggccttgatatttaatgctgggagatggaanaacttcgaatgctctct 60	
DB	74 AAGGAGGCTCAGGCTTCGATTTTAATAGCATGTGAGAAACATGGAAATGTTCTT 133	
OY	61 ggaatcacaagaagaactaagaagatccattccagaanaacactgaacagcagcaggagt 120	
DB	134 GGATGTACCAACAGCAACTGAAAGATCTTTCAAGAAATACCTGAAACACACAGATGACT 193	
OY	121 atctggtctctctctcggagacactggcgagagcaactctctctcccgctgctggtgt 180	
DB	194 ACCTAGCTTTGCTTCGGCGGCTCTGTCGACGACCTTCTCTCCCGGTGACTGCGGTGT 253	
OY	181 atgtcacaattcttggtggtggatcattgycaatgtccgtgtgctgtgactgtcgc 240	
DB	254 ATGCTCTGATTTTGTGTGGTGGGGGCTTGCGCAATCTCTGTGTGCTTGATGATCTTC 313	
OY	241 agcaccaggtcttgaagagcccaactactactccttcacgctgacgtgctctctgacc 300	
DB	314 GGCACACAGCGATGAGAGACCCACCAATTAATTAATTAATTAATTAATTAATTAATTA 373	
OY	301 tccgtgctcgtctcctcttggaatgcccttgaggtcattatagaatgtggtgcgaactac 360	
DB	374 TCTTACTGCTGCTTCTCGGGATGCCCTTGAGAGTCTATGATGTGTGGCCACTACCT 433	
OY	361 tctgtctcgggcgcgttggtgtctactcttaagaagcggcctcttgaagacgtgtgtc 420	
DB	434 TCTCTTTTGGGGCGGTGGGCTCTGCTTCAACACGGCCCTCTTTCGAGACGATGTGTT 493	
OY	421 cctgcattcctcagcatcacacacgctcaggtgtgagcgtcagctgacatcctcacaac 480	
DB	494 CCTCATCTCAGCTACACCACTGATGAGGTGAGAGGCTATGATGAGCGCTCTCACCG 553	
OY	481 tc 482	
DB	554 TC 555	
RESULT 2		
LOCUS	BE871983	
DEFINITION	601448032p1 NIH_MGC_65 Homo sapiens cDNA clone IMAGE:3852151 5',	
ACCESSION	BE871983	
VERSION	BE871983.1 GI:10320759	
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.	
TITLE	NIH-MGC http://mgc.ncl.nih.gov/	
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)	
	Unpublished (1999)	

COMMENT	CONTACT: ROBERT STRUNBERG, PH.D. EMAIL: CGAPBS-REMAIL.NHL.GOV TISSUE PROCUREMENT: ATCC CNA LIBRARY PREPARATION: LIFE TECHNOLOGIES, INC. CNA LIBRARY ARRAYED BY: THE I.M.A.G.E. CONSORTIUM (LLNL) DNA SEQUENCING BY: INCYTE GENOMICS, INC. CLONE DISTRIBUTION: MGC CLONE DISTRIBUTION INFORMATION CAN BE FOUND THROUGH THE I.M.A.G.E. CONSORTIUM/LLNL AT: http://image.llnl.gov PLATE: LLM9573 ROW: P COLUMN: 08 HIGH QUALITY SEQUENCE START: 11 HIGH QUALITY SEQUENCE STOP: 602. LOCATION/QUALIFIERS
FEATURES	1..929 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:3852151" /clone_1lb="NH_MGC_65" /tissue_type="adenocarcinoma" /lab_host="DH10B (phage-resistant)" /note="Organ: colon; Vector: pCMV-SPOK6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.8 kb. Library constructed by Life Technologies."
BASE COUNT	186 a 252 c 228 g 263 t
ORIGIN	
Query Match	26.1% Score 339; DB 11; Length 929; Best Local Similarity 86.1%; Pval: No. 5.7e-72; Matches 516; Conservative 0; Mismatched 65; Indels 18; Gaps 12.
1	aggagagcgcagcct-tgattctaagtcaaggatgaaacatcgaatgctcc 59
Db	107 AGGGAGGCTCAGGCGCTGTGATTTAATGTGAGGATGAAATTCGAATGCTTCC 166
QY	60 tggatctaccagcagaactagaaagatccatctcsgaaacacctgaacagcagagag 119
Db	167 TGGATCTACCAAG-AAGAACTACAAATCATTCATCCAAAACACCTGAMACGCCAGAGG 225
QY	120 tatctggcc-ttcctctcggagaccgcggcgagcagcaatctctccctccgltctgtgt 178
Db	226 TATCTGGCCTTCCCTCTGGGACCTCGGGGACCACTTCTCCGCCGTCTGTGGT 285
QY	179 gtatgtccaattttgt 238
Db	286 GTATGTGCCAATTATGTGGGGGGGTATGTGGGCAATGTCTGGTGTGATCTCT 345
QY	239 gaagacacaggtatgaagagcgcacacacacacacacacacacacacacacacacac 297
Db	346 GGAGCAACAGGCTATGAGACGCCCAACAA-TACTACCTCTTACGCTTGCGGCTCTCG 404
QY	238 acctctcgg--tcctgtctcttgaaatgccccttgagtgct--atgaatgttgcgcaa 352
Db	405 ACCTCTGGGCTCTGTGCTCTCTTTGGAGAGCCCGGTGAGAGCTAATGACATGTGGCGCA 464
QY	333 ctacccttctgttgcggccc--gttggctgtctaactcaagaagcgcccttttgagac 409
Db	465 CTACCCCTTTTGGGTGGGCGCCGTGTGGGCGTACATTCAGAGAGGCGCCCTTTTGGAGC 524
QY	410 cgtgtgcttcgctcacaatcagatcacacacgctcaagcgttggagcgctacgttgccat 469
Db	525 CGTGTGCTTGGCTCTCATCTCAGATCT-CACGCTCAG-GTGAGCGCTTACGTGGCAAT 582
QY	470 cctaacccgttcgcgcgccaacatgcagagcaccgcgcgcgcgcgccttcagaatcctcg 529
Db	583 CCTACACCCGTT-CGCGCCAAATTGAGGAGCACCGGGGTGCGGGGCTTCAGAAATCTCGG 641
QY	530 catgtctggaggtctctcgtgtcttcttcccttgcacacacacacacacacacacacac 588
Db	642 AATCTCTGGGCTACTCGAGCACTTCTCCCTG--CAAAACGACATCATGGGCTCA 698

RESULT	3
LOCUS	N45474
DEFINITION	yys9b04.r1 Soares_multiple_sclerosis_2NBHMSP Homo sapiens CDNA clone IMAGE:277807 5', mRNA sequence.
ACCESSION	N45474
VERSION	N45474.1 GI:1186640
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	1 (bases 1 to 414) Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaaskis,E., Waterston,R., Williamson,A., Woldmann,P. and Wilson,R. The Mashu-Merck EST Project Unpublished (1995)
TITLE	Contact: Wilson Rk
JOURNAL	Washington University School of Medicine 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810
COMMENT	Email: est@wustl.wustl.edu This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (infoimage.llnl.gov) for further information. Seq primer: T7 High quality sequence stop: 264. Location/Qualifiers 1..414 /organism="Homo sapiens" /db_xref="GDB:3886183" /db_xref="taxon:9606" /clone_image="IMAGE:277807" /clone_id="Soares_multiple_sclerosis_2NBHMSP" /sex="male" /tissue_type="multiple sclerosis lesions" /dev_host="Age 46" /lab_host="DH10B (ampicillin resistant)" /note="Vector: pT73D (pharmacia) with a modified polylinker V.type: phagemid; Site_1: Not I ; Site_2: Eco RI ; 1st strand cDNA was primed with a Not I - oligo(drf) primer [5', TGTTACCATCTGTGAAGTGCGGGCCGCATTCTTTTTTTTTTTT 3'] , double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M.Falima Bonaldo. RNA from 4 multiple sclerosis lesions from one patient was kindly provided by Dr. Kevin G. Becker (NINDS/NIH).
BASE COUNT	113 a 123 c 71 g 105 t 2 others
ORIGIN	
Query Match	23.3% Score 302 DB 11 Length 414;
Best Local Similarity	100.0%; Pred. NO. 4.7e-63;
Matches 302; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Oy	997 ccattatctataaacactgtctgcgcgcttcacagcagcatccagaagtgtaccttt 1056
Dd	1 ccattatctatTAACtACTGTCGTGCCGCTTCACAGCAGCATTCAGATGATCTCTT 60
Oy	1057 cttccacaacacagtgycgaacttcacgatggaccacagttgcacctgccagggaaca 1116
Dd	61 cttttccAAACAACAGGCACTCCACGATGACCCACAGTTGGCACCTGCCAGCGCAACA 120
Oy	1117 tctctcgacagaaatgccaactttggyggagcgcaagcaaatataggctcccataatccat 1176
Dd	121 tctttctGACAGATGCCACTTTGTGGAGACTGACCCGAAGAATATTAGGTCGCCAATTCCT 180

QY	1177	gtcagtcattccatgacacacitcctcaacccccaacagccctctctctgtatgaacagaatgtaaa	1236
Db	181	gtcagtcattccatgacacacitcctcaacccccaacagccctctctgtatgaacagaatgtaaa	240
QY	1237	gaacaaactatcaaacgtccacitcacttcaacaacagtgatcttcacagactgactcc	1296
Db	241	gaacaaactatcaaacgtccacitcacttcaacaacagtgatcttcacagactgactcc	300
QY	1297	tc 1298	
Db	301	tc 302	
RESULT	4		
LOCUS	BE751310	465 bp mRNA	EST 25-APR-2001
DEFINITION	203132 MARC 2B0V Bos taurus cDNA 5', mRNA sequence.		
ACCESSION	BE751310		
VERSION	BE751310.1	GI:10165239	
KEYWORDS	EST.		
SOURCE	COW.		
ORGANISM	Bos taurus		
REFERENCE	Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Cetartiodactyla: Ruminantia: Pecora: Bovidae: Bovidae: Bovinae: Bos.		
AUTHORS	Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T., Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G., Petten,G., Holt,I., Karanymcheva,S., Liang,F., Quackenbush,J. and Keefe,J.W.		
TITLE	Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle		
JOURNAL	Genome Res. 11 (4), 626-630 (2001)		
MEDLINE	21180013		
COMMENT	Contact: Smith TPL USDA, ARS, US Meat Animal Research Center PO Box 166, Clay Center, NE 68933-0166, USA Tel: 402 762 4366 Fax: 402 762 4390 Email: smith@email.marc.usda.gov Single pass sequencing. Bases called and alt_trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -minmatch 12 options. PCR primers FORWARD: AGGAACAGCATGACCAT BACKWARD: GTTTTCCAGTCACGAC Plate: 43 row: 0 column: 6 Seq primer: ATTTAGGTGACATATAG. Location/Qualifiers 1..465 /organism="Bos taurus" /db_xref="taxon.9913" /clone_lib="MARC 2B0V" /lisse_type="pooled" /lab_host="DH10B" /note="Vector: pCMV SPORT6; Site.1: XbaI; Site.2: XhoI; library made from pooled tissue from testis, thymus, semitendinosus muscle, longissimus muscle, pancreas, adrenal, and endometrium."		
FEATURES			
source			
BASE COUNT	88 a 129 c 113 g 135 t		
ORIGIN			
Query Match	23.1%	Score 299.2; DB 10; Length 465;	
Best Local Similarity	85.2%	Pred. No. 2.3e-62;	
Matches 334; Conservative	0; Mismatches 58; Indels 0; Gaps 0;		
QY	1	agggaagccctcagcgccttgatcttcaatcgtcaggaatggaacacatgcttcct 60	
Db	74	AGAGAGGCTCAGCGCTTGATCTTAAATAGCAGCATGTATGAAAAACATGAGATGTTTCTT 133	
QY	61	ggatctaccagcagaacatagatccattccagaaacacctggaacagcaccaggaagt 120	

Db 134 GATGTACCAAGAGAAAGATGATCTTTCAAGAAATACGTGAACACAGATACAT- 193
OY 121 atctgacccctccctcgcgcagccctcgcgcagccactctcccccgtctcgtggtc 180
Db 194 ACCGTACTTGGCTCTCGGGGCTCTGCGACGACCTCTTCCCGCTACGCGGCTG 253
OY 181 atgcgaactctctgctgctgagtcacatcgcagctcgtcgtcgtcgtcgtcgtc 240
Db 254 ATGCTCTGATTTTGTGCTGGGGCTGTTGGCATCTCCGTGTGTGTGTGTGTTC 313
OY 241 agcaccagcctcgtcgcgcagccacacactcactcctctcagcctcgtcgtcgtc 300
Db 314 GGCACCAAGAGATGAAGACCCACCAATATACCTCTTCAAGCTTGGCTGTCCGACC 373
OY 301 tctgtctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 360
Db 374 TCTGTACTTGTCTCTGCGGATGCTGGAAGCTATGATGTGGCGCACTACCTCC 433
OY 361 tctgtctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 392
Db 434 TCTGTCTTGGCGGCTGTGCTCTACTTCAAG 465

RESULT 5
LOCUS CNS036MH/c
DEFINITION Tetradon nigroviridis genome survey sequence T7 end of clone
21622 of library G from Tetradon nigroviridis, genomic survey
sequence.
ACCESSION AL230570.1 GI:7889565
VERSION AL230570.1
KEYWORDS GSS: genome survey sequence.
SOURCE Tetradon nigroviridis.
ORGANISM Tetradon nigroviridis.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodon.
1 (bases 1 to 941)
Roest-Crolius, H., Jallion, O., Dasilva, C., Fizes, C., Fisher, C.,
Bonneau, L., Billault, A., Quetier, F., Saurin, W., Bernot, A. and
Welsenbach, J.
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetradon nigroviridis
Unpublished
2 (bases 1 to 941)
Roest-Crolius, H., Jallion, O., Dasilva, C., Fizes, C., Fisher, C.,
Bernot, A., Fizes, C., Mincker, P., Brothier, P., Quetier, F.,
Saurin, W. and Welsenbach, J.
Human gene number estimate provided by genome wide analysis using
Tetradon nigroviridis DNA sequence
Unpublished
3 (bases 1 to 941)
Roest-Crolius, H., Jallion, O., Dasilva, C., Fizes, C., Fisher, C.,
Bernot, A., Fizes, C., Mincker, P., Brothier, P., Quetier, F.,
Saurin, W. and Welsenbach, J.
Human gene number estimate provided by genome wide analysis using
Tetradon nigroviridis DNA sequence
Unpublished

JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

FEATURES
source
1. 941
/organism="Tetradon nigroviridis"
/db_xref="taxon:99883"
/clone="21622"
/clone_lib="G"
/note="Genoscope sequence ID : COAG216DCL1P1-end : T7"
BASE COUNT
197 a 277 c 327 g 132 t 8 others

Query Match 16.3%; Score 211.2; DB 13; Length 941;

Best Local Similarity 67.3%; Pred. No. 7e-41;
Matches 325; Conservative 1; Mismatches 155; Indels 2; Gaps 2;
OY 108 agcaccagagatcctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 167
Db 568 AGGAGAGAGGCGACCTCTGCGAGGTACCGGGGCCCCCGAGTGGCGCTCTCTCC 509
OY 168 gttctcgttggctgctgctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 227
Db 508 GTCTGCTTCCCTTACCTCTGCGCATCTTCTGTGTGGGGTCTTGGGCAACTCTC 449
OY 228 ctgtgctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 287
Db 448 GCGGTATCTCTCGCTTCCGGGCGAGCGACCGCCACCAATATCTACTGTCTGA- 390
OY 288 gctgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 347
Db 389 GCGGCGTCCGACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 330
OY 348 cgcacactacactctctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 407
Db 329 CGGACTACCT 270
OY 408 accgtgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 467
Db 269 ACGGTGT-STTGGCTCTCATCTCTCAAGCTCAGCGCGCTCAGCGTGAAGCGCTAC 211
OY 468 atcctacacccgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 527
Db 210 GTGTGTGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 151
OY 528 ggcacgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 587
Db 150 CTCCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 91
OY 588 aag 590
Db 90 GAG 88

RESULT 6
LOCUS R13890
DEFINITION yf62a02.r1 Soares infant brain IN1B Homo sapiens cDNA clone
IMAGE:26587 5' similar to contains 11 repetitive element ;, mRNA
sequence.
ACCESSION R13890
VERSION R13890.1 GI:766966
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 450)
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman,
M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J.,
Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevaskis, E., Waterston,
R., Williamson, A., Wohlmann, P. and Wilson, R.
The WashU-Merck EST Project
Unpublished (1995)
Contact: Willson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
Insert Size: 1558
High quality sequence stops: 342 Source: IMAGE Consortium, LINT.
This clone is available royalty-free through LINT; contact the
IMAGE Consortium (info@image.lint.gov) for further information.
Insert Length: 1558 Std Error: 0.00
Seq primer: M13RPI
High quality sequence stop: 342.

JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Db	368	CGCCGCGTGCATCTGCTGGGTGCTCTCTGCGGCGCTTGGCCCTCCCATGGCGGGTATTC	367
0y	579	catggcatcaagttcca 595	
Db	368	ATGGGACAGAAAGACACGA 384	
RESULT 12			
CNS039RV	975 bp	DNA	GSS 15-MAY-2000
LOCUS			
DEFINITION	Tetraodon nigroviridis genome survey sequence T7 end of clone 008M03 of library G from Tetraodon nigroviridis, genomic survey sequence.		
ACCESSION	AL234292.1	GI:7893427	
VERSION	GSS: genome survey sequence.		
KEYWORDS	Tetraodon nigroviridis.		
SOURCE	Tetraodon nigroviridis		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorphi; Acanthopterygii; Percomorphi; Tetraodontiformes; Tetraodontidae; Tetraodon.		
REFERENCE	1 (bases 1 to 975) Roest-Crollius,H., Jalllon,O., Dasliya,C., Fizames,C., Fisher,C., Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and Weissenbach,J. Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis		
TITLE	Unpublished		
JOURNAL	2 (bases 1 to 975) Roest-Crollius,H., Jalllon,O., Dasliya,C., Bouneau,L., Fisher,C., Bernot,A., Fizames,C., Wincker,P., Brothier,P., Quetier,F., Saurin,W. and Weissenbach,J. Human gene number estimate provided by genome wide analysis using Tetraodon nigroviridis DNA sequence		
REFERENCE	Unpublished 3 (bases 1 to 975) Genoscope. Direct Submission Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases This sequence is a single read and was generated as part of a large scale clone-and-sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/tetraodon .		
FEATURES	Location/Qualifiers 1..975 /organism="Tetraodon nigroviridis" /db_xref="taxon:99883" /clone="008M03" /clone_1lb="G" /note="Genoscope sequence ID : COBG008AG02LP1-end : T7"		
SOURCE	164 a 332 c 295 g 179 t 5 others		
BASE COUNT			
ORIGIN			
Query Match	7.3%	Score 95.2;	DB 13; Length 975;
Best Local Similarity	54.0%;	Pred. No.1.2e-12;	
Matches 218; Conservative	0;	Mismatches 183;	Indels 3; Gaps 1;
0y	175	tgggtatgatgccaatttttggtggtgggggtcatatggcaatgtcctggtgtgctgtgga	234
Db	231	tgggtggcggcgtccctcagacctgtgtgacccgtatggcgaacatcctgcatgctgtcca	290
0y	235	tctctgacgacagcagcgtatgaagacgcccacacactactactcttcagccctgsgtct	294
Db	291	tcaagctacacccgacacacctgcagacagcgtcacaactactactctgtttcacacctgtgg	350
0y	295	ctgacctccctggtctcgtcctcttgaaatgcccctggaaggtctctatgagatgtgycgcaact	354
Db	351	ccgacctgattatcgtggcgtgctgtccatgaactgacacccgtatcattacatcgtgattggct	410
0y	355	acctctttgttctgggcccgtggcgtgctgctactcaagaagcgcctctttgagaccgtgt	414

[illegible]

DB 775 CATGACGATCCGGGAGGACCTACGGGCGCCGCTGACCTTCCCTTACCTGCGGCGGCG 716
 Oy 372 cccggtggtgtacttactcaagagcgcctcttgagacggtgtgtctgctccatcc 431
 Db 715 AACCTCTCTGTATACATMACACTTCTGTTCGAGGCGGGGACACACCCACGATGTGG 656
 Oy 432 agcatcacacacgctgagcgtgagcgtacgtgagcgtacgtacgttccggtccaa 491
 Db 655 DACATAGCCACCTTDCCTTCAGGCGCTACGTGGCCTTTCGCACCCCTTCGCTACAG 596
 Oy 492 ctgagagacacccggtgagcgtccctcagatcctgagcgtgtctggtccgtg 551
 Db 595 GCCCTGGG---CGGAGGCGGACACMACGCCCTTCGCTTGGCCCTKCGTGTGTCGGG 539
 Oy 552 cctctctccctgcccacacacgcatcgtgcatcgaagtcactactcc 605
 Db 538 CTGTGGCCCTGCGCTGCTCATCGCACGAGACGACAGNACATCCCTC 485

RESULT 14

CNS04P41 1079 bp DNA GSS 24-MAY-2000
 LOCUS Tetraodon nigroviridis genome survey sequence PUC-ORI end of clone
 DEFINITION 125F12 of library G from Tetraodon nigroviridis, genomic survey
 accession.
 AL300843.1 GI:8177862
 VERSION GSS: genome survey sequence.
 KEYWORDS Tetraodon nigroviridis.
 SOURCE Tetraodon nigroviridis
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorphi; Acanthopterygii; Percomorphi; Tetraodontiformes;
 Tetraodontidae; Tetraodon.
 1 (bases 1 to 1079)
 Roest-Crollius,H., Jallou,O., Daaliva,C., Fizames,C., Fisher,C.,
 Bouneau,L., Billault,A., Queller,F., Saurin,W., Bernot,A. and
 Weissbach,J.
 Characterization and repeat analysis of the compact genome of the
 freshwater pufferfish Tetraodon nigroviridis
 Unpublished
 2 (bases 1 to 1079)
 Roest-Crollius,H., Jallou,O., Daaliva,C., Bouneau,L., Fisher,C.,
 Bernot,A., Fizames,C., Winkler,P., Brotlier,P., Queller,F.,
 Saurin,W. and Weissbach,J.
 Human gene number estimate provided by genome wide analysis using
 Tetraodon nigroviridis DNA sequence
 Unpublished
 3 (bases 1 to 1079)
 Genome scope.

JOURNAL Direct Submission
 REFERENCE Submitted (12-APR-2000) to the EMBL/Genbank/DBJ databases
 TITLE This sequence is a single read and was generated as part of a large
 COMMENT scale clone-end sequencing project of the Tetraodon nigroviridis
 genome. For more information, please take a look at
 http://www.genoscope.cns.fr/Tetraodon.

FEATURES

source
 1..1079
 /organism="Tetraodon nigroviridis"
 /db_xref="taxon:99883"
 /clone="125F12"
 /clone.lib="G"
 /note="Genoscope sequence ID : C0BG125DC06SP1-end :
 PUC-ORI"

BASE COUNT 196 a 347 c 270 g 259 t 7 others
 ORIGIN

Query Match 73%; Score 94.6; DB 13; Length 1079;
 Best Local Similarity 64.7%; Pred. No. 1.7e-12;
 Matches 139; Conservative 1; Mismatches 75; Indels 0; Gaps 0;

Oy 839 tgccttgcttagtgcgttgcctatcgttgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 898
 Db 457 TGTCTGT 516
 Oy 899 cagcttctgtgagagagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtg 958
 Db 517 GAGCTGATGAGAGCTCTCTCTCCGACGCGACGTGAGGAGCTCTGAGAGGTCTACAGTAT 576
 Oy 959 gtcagtgcttct 1018
 Db 577 CTCTGAGAGTCT 636
 Oy 1019 tccgccttcacgagcagcttcacagatgtatct 1053
 Db 637 CACCCGCTTCAGGAAACCTTCAGCCACCTCACT 671

RESULT 15

AK004891 2611 bp mRNA HTC 05-JUL-2001
 LOCUS Mus musculus adult male liver cDNA, RIKEN full-length enriched
 DEFINITION library, clone:J300003P04, full insert sequence.
 AK004891
 ACCESSION AK004891.1 GI:12836417
 VERSION CAP trapper.
 KEYWORDS Mus musculus (strain:C57BL/6J) adult male liver cDNA to mRNA,
 clone.lib:RIKEN full-length enriched mouse cDNA library
 clone:J300003P04.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 2611)
 Carninci,P. and Hayashizaki,Y.
 High-efficiency full-length cDNA cloning
 Methods in enzymology. 303, 19-44 (1999)
 99279253
 MEDLINE 10349636
 PUBMED
 REFERENCE 2 (bases 1 to 2611)
 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
 Itoh,M., Kono,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 Genome research. 10 (10), 1617-1630 (2000)
 20499374
 JOURNAL
 MEDLINE 11042159
 PUBMED
 REFERENCE 3 (bases 1 to 2611)
 Shibata,K., Itoh,M., Aizawa,K., Nagaoaka,S., Sasaki,N., Carninci,P.,
 Kono,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M.,
 Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishino,T., Harada,A.,
 Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kasai,H.,
 Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watabiki,M.,
 Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J.,
 Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multichannel sequencer
 Genome research. 10 (11), 1757-1771 (2000)
 20530913
 JOURNAL
 MEDLINE 11076861
 PUBMED
 REFERENCE 4 (bases 1 to 2611)
 The RIKEN Genome Exploration Research Group Phase II Team and the
 FANTOM Consortium.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409, 685-690 (2001)
 5 (bases 1 to 2611)

Adachi,T., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A.,
 Arikawa,T., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M.,
 Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K., Hiraoka,T., Horii,F.,
 Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kato,H., Kawai,J.,
 Kojima,Y., Kono,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T.,
 Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y.,
 Okido,T., Owa,C., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H.,
 Sasaki,D., Shibata,K., Shibata,Y., Shingawa,A., Shikaki,T.,
 Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,

THIS PAGE BLANK (USPTO)

PUBMED
10899166
REFERENCE
2 (pages 1 to 1298)
AUTHORS
Bonati, J.A., Raddatz, R., Wilson, A. and Borowsky, B.
TITLE
Direct Submission
JOURNAL
Submitted (25-MAY-2000) Target Discovery and Assessment, Synaptic
Pharmaceutical Corporation, 215 College Road, Paramus, NJ 07652,
USA

FEATURES	location/Qualifiers
source	1..1298
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="5"
gene	map="cg31.1-q31.3"
	1..1298
	/gene="NMUR2"
CDS	27..1274

```

/gene="NMUR2"
/note="G protein-coupled receptor"
/codon_start=1
/product="neuromedin U receptor 2"
/protein_id="AAC24794.1"
/db_xref="GI:10946203"
/translation="MSGMEKIQANSAWYQOKLEDPFOKHLNSTEYLAFLDGPRRSH
FLPVSVPVPLFVGVIGNLVLCVLIQOHAMKRPNTYLYSLASDLVLLDMPLE
VEEMRWYRPELFEGVGFYKALFEVCEIISITTVSERYAAILHPFAKQSTR
RRALRTIGIWGSEVLSRPTSEIHGKHFHPGSLVPGSACTCTVIKPMVNIQO
VTSGLFLLPRTVLSVLYLMALRLKDKDSLEADGNANDRPCRSYNNKLFVLV
FALCAAFPHIDRLFSEVEEWSLSLAVFNLVHVSQVFFLSAVNDPIYNLSRR
QAACQNYISFHHQMSHDQPLPPAQRNIFLTCHEVELTEDIGPQFQSSMNSH
LPTALSSQMSRTVYOSHFNRKT"

```

Query Match	100.0%;	Score 1298;	DB 9;	Length 1298;
Best Local Similarity	100.0%;	Pred. No. 3.6e-259;		
Matches 1298;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

OY	1	agggagggctcagagccttggattcttaagtcgaggaatggaanaacttcagaatgtctcc	60
Db	1	AGGGAGGGCTCAGGCCTTGGAATTTAAATGTCAGGATGGAAAAACTTCAGATGCTTCC	60
OY	61	ggatctaccagcagaagaactagaagaatccattccagaabaaccttgaaacagcacgagagtg	120
Db	61	GGATCTACCACACAGAAACTAGAAAGATTCATTCCAAACAAACACTGTAAACAGCAGGAGAGT	120
OY	121	atctgagcttcctctgaggaacctggcgagcgaaccttcctctcccgctgactgtgagtgat	180
Db	121	ATCTGAGCTTCTCTCTGGCGGAGCTGGGGCAGCACAATTCTCTCCCGGTGTGTGGGT	180
OY	181	atgtgcaaatcttctgtgtggtgggtcaattggaatgtccctgtgtcctgtgtaattctgc	240
Db	181	ATGTGCAAATTTTGTGTGGTGGGTCATTTGGCAATGTCTGTGTGCTGTGGGTGATTTGTC	240
OY	241	agagccagagcatgaaagacgcccacaacactactacctcttcagacctggcaggtctctgac	300
Db	241	AGACACAGAGCATATGAAGAGGCCACACAACTACTACTCTTCAAGCTTGGGGTCTTGAC	300
OY	301	tccctgtccctgcctcctcttgaaatgccccctggaggtctataagagatctgtgcgcaactaccct	360
Db	301	TCCCTGTCCTGCTCCTTGTGAATGCCCTTGAGAGGTATAGATGTGGGCAACATCACTT	360
OY	361	tctgttctggagccgctgtggctgtcctactccaagaacgagccctcttgaagaacgtgtgctcgg	420
Db	361	TCTTGTTTGGGCCCCCTGGGCTGTCACTTCAACACGGCCCTTTTAAAGACCGTGTCTTTCG	420
OY	421	ccctcaatcccaagacatcaacaacgctcagcgctgaaagcgtacgtgtggcatctcaaacccgt	480
Db	421	CCTCATCTCTCAGCATCAACACACACGTCACGCTGAGACGCTACGTGGCATCTCAACACCGT	480
OY	481	tccgcgcacaactctcagagcaccccgcgccgagccctcaagatcctctgcatcgtctggg	540
Db	481	TCCGCGCCAACTCTGAGAGCACCCCGGCCCTCAGATCTCTCGGCATCTGCTGGG	540

QY	541	gcctctccgagctctctctccctccgccaacaccagcaatcccaatggagatcaagttcccaact	600
Db	541	gcttctccgctctctctctccctccgccaacaccagatccatggatcnaagttcccaact	600
QY	601	tccccaatggagtcctctggtctccaggtctcggccactctgacgtgcatcaagcccaatggaga	660
Db	601	tcaccaatggagtcctctggtctccaggtctcggccactctgacgtgcatcaagcccaatggaga	660
QY	661	tctacaattcatcaatccaggtcgaacctctctctactatctacacctctcccactgactgca	720
Db	661	tcctacaattcatcaatccaggtcgaacctctctctactatctacacctctcccactgactgca	720
QY	721	tcagatgtctcttacttactctcaatggacatccagaactcaaaaagaagaactcttgaaggcag	780
Db	721	tcagatgtctcttacttactctcaatggacatccagaactcaaaaagaagaactcttgaaggcag	780
QY	781	atgaagaaggaatgcacaatatctcaaaagccctcagagaataatcgatcaacaagaatgtctgtt	840
Db	781	atgaagaaggaatgcacaatatctcaaaagccctcagagaataatcgatcaacaagaatgtctgtt	840
QY	841	tcctgtctcttagtgltgtgctatctgtgttggcccgcttcacacatgacgcgactctcttca	900
Db	841	tcctgtctcttagtgltgtgctatctgtgttggcccgcttcacacatgacgcgactctcttca	900
QY	901	gcttctgtggaagagtgagagtgaaatccctcgctctgtctgtgtgttcaacctgtgcaatgtgt	960
Db	901	gcttctgtggaagagtgagagtgaaatccctcgctctgtgtgtgttcaacctgtgtgtgt	960
QY	961	caggctgtctcttctcaactgaagctcgaagctgtgcacaaccccatatctataacctactgtctc	1020
Db	961	caggctgtctcttctcaactgaagctcgaagctgtgtgcacaaccccatatctataacctactgtctc	1020
QY	1021	gccgcgttcagagcagcatcttcagaatgtgactctctcttccacaagaacagtgtgcaactccc	1080
Db	1021	gccgcgttcagagcagcatcttcagaatgtgactctctcttccacaagaacagtgtgcaactccc	1080
QY	1081	agcatgatcccaacagttgcacacctgcgcagcgagaaatctctctcgaagaaatgagccaattg	1140
Db	1081	agcatgatcccaacagttgcacacctgcgcagcgagaaatctctctcgaagaaatgagccaattg	1140
QY	1141	tggagctgaacgaagaataatagttccccaattcccaatgacgtgcatgcatcgaacacacttc	1200
Db	1141	tggagctgaacgaagaataatagttccccaattcccaatgacgtgcatcgaacacacttc	1200
QY	1201	acctcccaacagccctctctagtgaaacagatgtgcaagaacaacatatacgaagttccaact	1260
Db	1201	acctcccaacagccctctctagtgaaacagatgtgcaagaacaacatatacgaagttccaact	1260
QY	1261	tttaacaaaacctgaattctcttcagagcgtcgaactctctc	1298
Db	1261	tttaacaaaacctgaattctcttcagagcgtcgaactctctc	1298

RESULT	2		
AXI09234			
LOCUS	AXI09234	1658 bp	DNA
DEFINITION	Sequence 1 from Patent WO0125269.		
ACCESSION	AXI09234		
VERSION	AXI09234.1	GI:13924107	
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi		
	Mammalia; Eultheria; Primates; Carnivora; Homini; Hominidae; Homo.		
REFERENCE	1 (bases 1 to 1658)		
AUTHORS	Deleersnyder, W., Berger, C., Loeken, C., Nys, G., and Venema, J.		
TITLE	Human g-protein coupled receptor		
JOURNAL	Patent: WO 0125269-A 1 12-APR-2001;		
	Solvay Pharmaceuticals B.V. (NL)		
FEATURES	Location/Qualifiers		
source	1..1658		
	/organism="Homo sapiens"		

CDS
/db_xref="taxon:9606"
55..1302
/note="IGS4 long version"
/codon_start=1
/protein_id="CAC3777.1"
/db_xref="GI:13924108"
/translation="MSGMKLQNSMIVYQKLEDFQKHLNSTEYLAFLGPRRSHF
FLPVSVVPIFEVGVIGNVLVCLVLIQHOAMKPTNYLFSLAUSDLLVLLGMPLE
VYEMRNPLFEGVGYEKTALFEVCFASLITTVSVERVYAILHPRAKLOSTR
RRALRLDITVMGFSVLSLPNTSHIGIKFHFHPNSLVPGSATCVIKPMIYNITIQ
VTSLEFLIPMTIVLSVLYLMALRLKKSLEADGNANIQPKRSYKMKLEVLV
FALCMAPHIDRLFEFSFEWSESLAANFNLVHVSQVFLSSAVNPIIYNLSRR
QAFQOMVSSFHQMSQHDLPOLPPAQRNIFLFEVELTEDIGPFCOSSMNSH
LPTALSSEOMSRNTYQSFHNKT"
BASE COUNT 377 a 490 c 346 g 445 t
ORIGIN

Query Match 100.0%; Score 1298; DB 6; Length 1658;
Best Local Similarity 100.0%; Pred. No. 3,5e-259;
Matches 1298; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 agggagagctcagccttgatcttaatgctcagggatggaataacttcagaatgcttcc 60
DB 29 AGGGAGAGCTCAGCGCTTGATTTTATGTCAAGGATGAAAAACTTCAGAAATGCTTCT 88
QY 61 ggaatcacagcagaagaactgaatgccattccagaataacttgaaacagccaggaagt 120
DB 89 GGATCTACAGAGAGAAACTGAAGATTCATCCAGAAACCTGAAACACCGAGAGAGT 148
QY 121 atctggcctctctcggagagcctcggcagcgaactctctcccgctgctctgctg 180
DB 149 ATCTGGCCTCTCTCGCGAGACCTCGCGGACGACACTTCTTCTCCCGCTGCTGTGT 208
QY 181 atgtgcgaattttgt 240
DB 209 ATGTGCGAATTTTGT 268
QY 241 agcacaagagctatgaagaagcccaactactactctctcagcctggcggtctcgacc 300
DB 269 AGCACCAGAGCTATGAAGAGCCGACCACTACTACTCTTCAAGCCGCGGCTCTGACC 328
QY 301 tccctgctcctcctcttgaaatgagccttgagagctatgagatgctgagcgaactacc 360
DB 329 TCTTGCTCTGCTCTCTTGGAATGCCCTTGAGAGTCTATAGATGTGCCCACTACCTT 388
QY 361 tctgtctcggcgcgtggtgctgctacttcagaagcgcctctcttgagccgtgctgct 420
DB 389 TCTTGCTTGGGCGGCTGCTGCTCTACTTCAAGACGCGCTCTTGAAGACCGTGTGCTG 448
QY 421 cctcacaactcagatcacaccgctcagctgagagctcagctgagcgaactcaccgct 480
DB 449 CCTTCATCTCAGCATCACACCGTGAAGCTGACGTGCGCATCTACACCGCT 508
QY 481 tccgagcacaatgcaagcagccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 540
DB 509 TCCGCGCCCAATGTCAGACACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 568
QY 541 gctctcgcgtgctctctcctcctcgcgaacacagcaatccatgcaatgaagctccact 600
DB 569 GCTTCTCCGCTCTCTCTCTGCGCAACACAGCATGCGATGCGATCAAGTTCCACTACT 628
QY 601 tccccaatgagctcctcgtgctcagagctgagcgaactgtaacgaagcccaatgta 660
DB 629 TCCCAATGAGTCCGCTGCTGCGCAAGTTCGCAAGGTCATCAAGCCCATGTGGA 688
QY 661 tctcaatctcaatcagcagctcctcctcctatctcaactcctccccaatgagctgca 720
DB 689 TCTACATTTTCAATCAGGTCACCTCTTCTATTTCTACCTCTCCCATGACTGTCA 748
QY 721 tcaagtgcctctaactcaatcagcagctcagaactaaagaagaactctctgagcag 780
DB 749 TCAGTGTCTCTACTACTATGAGCTGAGACTAAAGAAAGCAATCTTGAAGCAG 808

QY 781 atgaaggagatgcaaatatctcaagaagcctctgaggaataatcagtcacaagatgctgtg 840
DB 809 ATGAGGAGATGCAATATTTCAAAGACCTCGCAAAATCAGTCAACAAATGCTTTTG 868
QY 841 tcttgctcctgctgtctgtatctgtctgagcccgcttcacatctgacgactctctca 900
DB 869 TCTTGCTCTTAGTGTGTGTATCTGTGGGCGCGCTTCCACATTCAGCCGACTTCTTCA 928
QY 901 gcttgctgagagagtgagtgaaatccctgctgctgctgtctcaactcgtccatgctgt 960
DB 929 GCTTGTGAGGAGTGAGTGAATCCCTGCTGCTGTGTTCAACCTCGTCAATGTGTGT 988
QY 961 caggtctctctctcctcagcagctcagctcagctcagcagcagcagcagcagcagc 1020
DB 989 CAGGCTCTTCTTCTACTGAGCTCAGCTGTCACCCCATATCTATCACTACTGTCTTC 1048
QY 1021 gccgcttcagagcagatcaccagaatgctatctctcttcacacaagaagctgcactcc 1080
DB 1049 GCGGCTTCAGGAGCATTCAGAAATGTATCTTCTTCCACAAACAGTGCACTTCC 1108
QY 1081 agcatgacccacagcttgcaacctgcccagcggaacatctccttgacagaatgccacttg 1140
DB 1109 AGCATGACCCACAGTTCGCCACCTGCCCAGCAATCTTCTCGACAGATGCACTTGG 1168
QY 1141 tggagctgacgagaatagatagctcccaatctcccatgctacatccatgcaactctc 1200
DB 1169 TGGAGCTGACCGAAGATATAGTCCCAATTCCTCATGTATCCATGCAACACTCTC 1228
QY 1201 acctccacaagcgcctctcagtgaaacagatgcaagaacaaactcaagaactcact 1260
DB 1229 ACTTCCCAACAGCCTCTCTAGGAACAGATGTCAGAAACAACTATCAAGCTTCTCA 1288
QY 1261 ctacaagaactgaaatctcttcagagctgactctc 1298
DB 1289 TTAACAAAACCTGAATCTTTCAGAGCTGACTCTCTC 1326

RESULT 3
AX109236 1658 bp DNA PAT 30-APR-2001
LOCUS
DEFINITION Sequence 3 from Patent WO0125269.
ACCESSION AX109236
VERSION AX109236.1 GI:13924109
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 1658)
Deleersnijder, W., Berger, C., Loeken, C., Nys, G. and Venema, J.
Human g-protein coupled receptor
patent: WO 0125269-A 3 12-Apr-2001;
Solvay-Pharmaceuticals B.V. (NL)
FEATURES
source location/Qualifiers
1..1658
/organism="Homo sapiens"
/db_xref="taxon:9606"
64..1302
/note="IGS4 short version"
/codon_start=1
/protein_id="CAC3777.1"
/db_xref="GI:13924110"
/translation="MEKLNMSWIVYQKLEDFQKHLNSTEYLAFLGPRRSHF
VSVVVPPIFEVGVIGNVLVCLVLIQHOAMKPTNYLFSLAUSDLLVLLGMPLE
VYEMRNPLFEGVGYEKTALFEVCFASLITTVSVERVYAILHPRAKLOSTR
RRALRLDITVMGFSVLSLPNTSHIGIKFHFHPNSLVPGSATCVIKPMIYNITIQ
VTSLEFLIPMTIVLSVLYLMALRLKKSLEADGNANIQPKRSYKMKLEVLV
FALCMAPHIDRLFEFSFEWSESLAANFNLVHVSQVFLSSAVNPIIYNLSRRQ
A FQOMVSSFHQMSQHDLPOLPPAQRNIFLFEVELTEDIGPFCOSSMNSH
LPTALSSEOMSRNTYQSFHNKT"
BASE COUNT 377 a 490 c 346 g 445 t
ORIGIN

Query Match 100.0%; Score 1298; DB 6; Length 1658;
 Best Local Similarity 100.0%; Pred. No. 3.5e-259;
 Matches 1298; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 agggagagctcagagccttgatlttaatgctcaggatggaagaaacttcagaatgctcct 60
DB 29 AGGGAGAGCTCAGAGCCTTGATTTAATGTCAGAGATGGAATAACTTCAGAAATGCTTCT 88
OY 61 ggaatcacaagcagaactagaatgcaatcctcagaagaacactgaaacagcaggagat 120
DB 89 GGATTCACAGCAGAAACTAGAAGATCCATTCAGAAACCCAGAACAGCAGCAGAGAGT 148
OY 121 atctggccttcctctgagagcctcgagcagcaactctctcctcccgctgctggtgtg 180
DB 149 ATCTGGCCTTCTCTGCGGACCTCGGGCGACACCTTCTTCTCCCGGTGTGTGTGTGT 208
OY 181 atgtgcaaatcttctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 240
DB 209 ATGTGCAATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 268
OY 241 agcagcagagctatagaagagcccaactactactctctcagcctgagcggtctctgac 300
DB 269 AGCAGCAGGCTATGAGAGCGCCACCACTACTACTCTTACACCTGTGGCGTCTGACC 328
OY 301 tctgtgtctgtcctcttggaatgccccctgagaggtctctagagatgtgagcgaactac 360
DB 329 TCTGTGCTCTGCTCTTGTGGAATGCCCTGTGAGGTCTATGAGATGTGTGGCGAATCT 388
OY 361 tctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 420
DB 389 TCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 448
OY 421 ccttcaatctcagatcacaacacgctcagagtgagagtgagtgagtgagtgagtgagtg 480
DB 449 CCTCAATCTCAGCATCACCATCGTACGCTGAGCGTGAAGCGTACGTCATCTACACCG 508
OY 481 tccgcagcaaaatgcagagcaacccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 540
DB 509 TCCGCCCAAACTGACAGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 568
OY 541 gctctcgtgtcctctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 600
DB 569 GCTTCCCGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 628
OY 601 tccccaatgggtcctcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 660
DB 629 TCCCAATGGGTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 688
OY 661 tctacaattcatalcagcaggtcactcctctcctcctcctcctcctcctcctcctcctc 720
DB 689 TCTACAAATTTCAATCATCGAGGTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 748
OY 721 tcaagtgtcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 780
DB 749 TCAGTGTCTCTACTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 808
OY 781 atgaaaggaatgcaaatcattcagaagacccctgagagaatcaatcagcacaagaatgtgt 840
DB 809 ATGAAAGGAATGCAAAATTTCAAGACCTCTGAGAAATAAGTCAACAGATCTCTGTG 868
OY 841 tcttggctctagtgtctgtatctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 900
DB 869 TCTTGTGCTTGTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 928
OY 901 gcttgtgtgagagatgagatgaaatccctgagcgtgtgtgtgtgtgtgtgtgtgtgtgt 960
DB 929 GCTTGTGTGAGAGATGAGATGAAATCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 988
OY 961 cagggtgtcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 1020
DB 989 CAGGTGTCTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1048

```

```

OY 1021 gcccttccagcagcatcaccagaatgtgactctctctcttccacaacagtggaactccc 1080
DB 1049 GCCGCTTCCAGGACGATTCAGATGTGATCTCTTCTTCCACMAACAGTGGCACTGCC 1108
OY 1081 agcatgaccacagatgtgccactcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1140
DB 1109 AGCATGACCCACAGTGTGCCACTGTCCAGCGGAACATCTTCTCTGACAGAAATGCC 1168
OY 1141 tgaagctgacaggaatagatggttccccaattcccaatgtaagtcataatgcaactctc 1200
DB 1169 TGAAGCTGACCGAAGATATAGTGTCCCAATTTCCATGTGATCATGCAACACTCTC 1228
OY 1201 acctcccaagcgcctctcctagtgaaacagatgcaagaacaaactcaaaagcttccact 1260
DB 1229 ACCTCCCAAGCGCCTCTCTAGTGAACAGATGTCAAGAACAAACTTCAAACTTCC 1288
OY 1261 ttaacaaaactgaaatctcttcagagctgaactctc 1298
DB 1289 TTAACAAAACCTGAATTTCTTTCAGAGCTGACTCTCTC 1326

RESULT 4
AX109238 1658 bp DNA PAT 30-Apr-2001
LOCUS AX109238
DEFINITION Sequence 5 from Patent WO0125269.
ACCESSION AX109238
VERSION AX109238.1 GI:13924111
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1658)
AUTHORS Deleersnijder, W., Berger, C., Loeken, C., Nys, G. and Venema, J.
TITLE Human g-protein coupled receptor
JOURNAL Patent: WO 0125269-A 5 12-Apr-2001;
Solvay Pharmaceuticals B.V. (NL)
FEATURES
source location/Qualifiers
1..1658
CDS
1658
/organism="Homo sapiens"
/db_xref="taxon:9606"
55..1302
/note="IGS4B long version"
/codon_start=1
/protein_id="CAC37773.1"
/db_xref="GI:13924112"
/translation="MSGMKELQNASHTYQKLEDPQKHLNSTERYLAFLCGRNSHF
FLPVSVVYPIFVGVIGNVLCVLVLOHQAARKTPNYLFLSLAVSDLVLLGMPLE
VYEMRNYPFLGPGVGCYKFTALFETVCPASILSTITVSEVRVAIHLPRFKLQSTR
RAALRLIGVMEFSVLSLPNTSINGIKFHPYPSLVGSAATCTVYIKPMYINFLIQ
VNSFLYVLPMTVISLYLALRLKDKSPADGSMANIORPKRSVYKMLFVLV
FAICNAPRHIDLFESFVEEMETLSLAAYFNVLHVYGVLYFYSVAVNPITVYMLSRF
QAAFOVISFHKHMSQHDPLPPRAQRIFLFECHFEVLIEDIGPFLCGSSVYNSH
LPTALSSSEMSSTNFQSFHFKT"
BASE COUNT 378 a 489 c 346 g 445 t
ORIGIN

```

Query Match 99.5%; Score 1291.6; DB 6; Length 1658;
 Best Local Similarity 99.7%; Pred. No. 7.4e-258;
 Matches 1294; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

OY 1 agggagagctcagagccttgatlttaatgctcaggatggaagaaacttcagaatgctcct 60
DB 29 AGGGAGAGCTCAGAGCCTTGATTTAATGTCAGAGATGGAATAACTTCAGAAATGCTTCT 88
OY 61 ggaatcacaagcagaactagaatgcaatcctcagaagaacactgaaacagcaggagat 120
DB 89 GGATTCACAGCAGAAACTAGAAGATCCATTCAGAAACCCAGAACAGCAGCAGAGAGT 148
OY 121 atctggccttcctctgagagcctcgagcagcaactctctcctcccgctgctggtgtg 180

```



```
Db 149 ATCTGGCCTTCCTCTGCGGAGCCTCGGCGAGCCACTTCCTCCCGTGTCTGTGTGT 208
Qy 181 ATCTGCGCAATCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 240
Db 209 ATCTGCGCAATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 268
Qy 241 AGCAACGAGGCTATGAAAGACGCGCCCACTACTACTCTCTCAAGCCTGTGCTCTGACC 300
Db 269 AACACAGAGGCTATGAAAGACGCGCCCACTACTACTCTCTCAAGCCTGTGCTGTGACC 328
Qy 301 TCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 360
Db 329 TCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 388
Qy 361 TCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 420
Db 389 TCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 448
Qy 421 CCTCCATCTCTAGACATCAACACCGTCAAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGC 480
Db 449 CCTCCATCTCTAGACATCAACACCGTCAAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGC 508
Qy 481 TCCGCGCAAACTGTGAGAGACCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540
Db 509 TCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 568
Qy 541 GCTTCTCGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 600
Db 569 GCTTCTCGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 628
Qy 601 TCCCGCAATGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 660
Db 629 TCCCGCAATGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 688
Qy 661 TCTACCAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 720
Db 689 TCTACCAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 748
Qy 721 TCAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 780
Db 749 TCAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 808
Qy 781 ATGAAAGGAAATGCAAAATATTTCAAAAGACCTGCAAGAAATGCAAGAAATGCAAGTGT 840
Db 809 ATGAAAGGAAATGCAAAATATTTCAAAAGACCTGCAAGAAATGCAAGAAATGCAAGTGT 868
Qy 841 TCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 900
Db 869 TCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 928
Qy 901 GCTTGTGTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 960
Db 929 GCTTGTGTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 988
Qy 961 CAGGTGTCTTATTTCTATCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 1020
Db 989 CAGGTGTCTTATTTCTATCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 1048
Qy 1021 GCGCGCTCTGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1080
Db 1049 GCGCGCTCTGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1108
Qy 1081 AGCAATGACCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1140
Db 1109 AGCAATGACCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1168
Qy 1141 TGGAGCTGACGAAAGTATAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 1200
Db 1169 TGGAGCTGACGAAAGTATAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 1228
Qy 1201 AACTGCGCAAGCGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1260
Db 1229 AACTGCGCAAGCGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1288
```

```
Qy 1261 ttacaaacccttgatctcttcacagagctgactctcc 1298
Db 1289 TTACAAAAACCTGATCTTTACAGAGCTGACTCTCC 1326

RESULT 5
AX109240
LOCUS AX109240 1658 bp DNA PAT 30-APR-2001
DEFINITION Sequence 7 from Patent WO0125269.
ACCESSION AX109240
VERSION AX109240.1 GI:13924113
KEYWORDS
SOURCE
human.
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 1658)
AUTHORS Deleersnijder,W., Berger,C., Loeken,C., Nys,G. and Venema,J.
TITLE Human g-protein coupled receptor
JOURNAL Patent: WO 0125269-A 7 12-APR-2001;
Solvay Pharmaceuticals B.V. (NL)
FEATURES
Location/Qualifiers
source 1..1658
/organism="Homo sapiens"
/db_xref="taxon:9606"
64..1302
/note="TGS4B short version"
/codon_start=1
/protein_id="CAC37774.1"
/db_xref="GI:13924114"
/translacion="MEKLOMASWYIOOKLEDDPOKHLNSTEYLAFLCGPRSHFLP
VSVVYPIFVGVGVIGVILVCLVILIOHQAAMKTPNLYFLSLVSLDLVILLPLLEYE
MRRNRYFELGPGCYRKTALFETVCASLSTITTVSEYVAILHFRKLOSTRRA
LRLGLVWGFVSLPLNTSHQIKFHPNSLVGSAFCTVIRKMYNPLIOVTS
FLFYLPLMTVISLVLYLMAIRLKRKDSLADBDGNANIORCRKSVKMLFVLVFAI
CHAPFIDILFEFSEWETESLAAVNLVAVGVLFYLSAVNPILVNLISRFQAA
FONVSSFHKOMHSOHPOLPPAQRNIFLTECHFVELFDIGFQLOSSVHNSHLP
ALSSEMSRFTNQSFFHFKT"
BASE COUNT 378 a 489 c 346 g 445 t
ORIGIN

Query Match 99.5%; Score 1291.6; DB 6; Length 1658;
Best Local Similarity 99.7%; Pred. No. 7 4e-258;
Matches 1294; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 aggggaagcctcaagccttgatctttaaagtcaggaatgaaactcaagaatgctcct 60
Db 29 AGGGGAGGCTCAGGCGCTTGATTTTAAATGTCAGGAGTGAATAAATCAGAAATGCTTCT 88
Qy 61 ggaatcaccgcaagaactgaagaatcattccaagaaccccgaaagacagcgaagagt 120
Db 89 GGAATCTACAGCAAGAACTGAAAGATCTCATTCAGAAACCTGAAACGACGAGAGT 148
Qy 121 atctgagcctccctctcgagaccccgagcagcaactctccctcccgctgtctgtggt 180
Db 149 ATCTGGCTTCTCTCTGCGGACCTCGGCGACGACACTTCTCTCCCGTGTCTGTGTGT 208
Qy 181 atgtgccaatcttctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 240
Db 209 ATGTGCCAATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 268
Qy 241 aggaagagctatgaagaagcccaaccaactaactcctcaagcctgtgagctctgacc 300
Db 269 AGCAACAGGCTATGAAACCCCAACACTACTCTTCTACGCTGTGCGGTCTGTGACC 328
Qy 301 tctgtgtcctctccctctgtgaatgcccttgagggttcta tgaatgtgtgtgtgtgtgtgt 360
Db 329 TCTGTGTCTGTCTCTTGTGGAATGCCCTTGAGAGCTATGAGATGTGCGGCAACCTT 388
Qy 361 tctgtgtcggccgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 420
```

Db	389	TCCTGTGGGGCCGCTGGGCTCTACTTCAACAGCGCCCTCTTTGAGACCGTGTGCTTCG	448
Oy	421	ccctcacccctcagaatcacaaccgctcagcgtgtagcgctacgttgccacatccataccggt	480
Db	449	CCTTCATCTCTCAGCATACACACGCTCAGGCTGAGACGGCTACCTGGCCATCTTACACCCGT	508
Oy	481	tcgcgcgaacaactgtaagagcaaccggcgccgggacctcaagatctcttgacatcgctctgg	540
Db	509	TCGCGGCCCAATGCAAGACACCCGCGCGCGGCGCTTCAGATTCCTGGCACTGCTGGG	568
Oy	541	gctctccgctgctctctccctcgcccaaacacagacatccatggaatcgaagttccactact	600
Db	569	GCTTCTCGGTGCTCTTCTTCCCTGCCCAACACGACATCCATGGCATCAAGTTCCACTACT	628
Oy	601	tcaccaatgagtgccctggtgccgaagtctcgccacctgtacggtcaatcacaagcccatgtga	660
Db	629	TCGCCCAATGGGCTCCTGGGTCGACGAGTGGGCACTGTACGGTCAATCAAGCCCATGTGA	688
Oy	661	ctctacaattctcatatcccaagtcacacctctctctatctactctctcccatgactgtca	720
Db	689	TCCTACAAATTTCTCATCTCCAGGTCACCTCTCTTCTTCTACTCTCTCCCCATGACTGTCA	748
Oy	721	tcagctgcctctcaactcaaccatcggagacatcaagaagaacaatatctcttgaggag	780
Db	749	TCAGTGCTCTCTACTACTCTCATGCACTCAGACTAAGAAACAAATCTCTTGAGGGAG	808
Oy	781	atgaagaggaaatgcaaatatcacaagaaccctgacagaataatcagtcacaagaatgacgtt	840
Db	809	ATGAAGGGGAATGCAATATTTCAAGCCCTGCAGAAATCACTCAACAAAGATGCTGTTG	868
Oy	841	tcctggtcttagtggtttgctatcgttgggcccggttccaatgaacgactctcttca	900
Db	869	TCCTTGCTTTACTAGTGTGGTATCTGTGGGCCCTCCCTCACATTTGACCCACTCTTCTTCA	928
Oy	901	gctttgtgaggagatggaatgaaatccctgcgcgcgtgtttaaactcgtgcaatgtgaggt	960
Db	929	GCTTTGTGGAGAGTGGAGTGAATCCCTGGCTGTGTGTCAACCTCGTCATGTGGTGT	988
Oy	961	cagtgctctctctcaaccctgagctcagctcgtataaccccatactataaccctactgtcc	1020
Db	989	CAGGCTGTATTATTACCTGAGACTCAGCTGTCAACCCATTATCTATACCTACTGTCTC	1048
Oy	1021	gccgcttccaggcagcatctcagaatgtaatctctctcttcacaacaagatggaactccc	1080
Db	1049	GCCGCTTCCAGGACACATCTCCAGAAATGTGATCTTCTTCCACAACAGTGGCACTCCC	1108
Oy	1081	agcagagaccbaagatgtgcaccctgcccagcggaaatatctctctgacgaatgaccaattg	1140
Db	1109	AGCATGATCCCAAGATGGCCACTCTGCCACGGGAACATCTTCTGTGACAGATCTCCACTTGG	1168
Oy	1141	tggagctgacggaaatataagttccccaattcccaatgtaagcatcccatgacaactctc	1200
Db	1169	TGGACTGACCGAATATATAGTCTCCCAATTTCTTATGTACGTCACTCCGTGCAAACTCTC	1228
Oy	1201	accctcccaacagccctctctctagtgaaacgaatgltcaagaaacaatatcaaaagcttccact	1260
Db	1229	ACCTCCCAACAGCCCTCTCTACTGAAACAGATGTCAAGAACAAATCATCAAACTTCCACT	1288
Oy	1261	ttaaacaaaacctgaaatctcttcaagagctgacactctctc	1298
Db	1289	TTTAAACAAAACCTGAAATTTCTTTGAGCGTGACATCTCTCTC	1326

RESULT	6
AB041228	
LOCUS	1248 bp mRNA
DEFINITION	Homo sapiens mRNA for G protein-coupled receptor TGR-1, complete cds.
ACCESSION	AB041228
VERSION	AB041228.1 GI:10257380
KEYWORDS	G protein-coupled receptor TGR-1.
SOURCE	Homo sapiens CDNA to mRNA.
ORGANISM	Homo sapiens
	21-SEP-2000

REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE	Hosoya,M., Moriya,T., Kawamata,Y., Ohkubo,S., Fujii,R., Matsui,H., Shintani,Y., Fukusumi,S., Habata,Y., Hinuma,S., Onda,H., Nishimura,O. and Fujio,M. Identification and Functional Characterization of a Novel Subtype of Neutrophin U Receptor
JOURNAL	J. Biol. Chem. 275 (38), 29528-29532 (2000)
PUBMED	10887190
RECORDS	2 (bases 1 to 1248)
AUTHORS	Shintani,Y., Moriya,T., Ohkubo,S. and Matsui,H.
TITLE	Direct Submission
JOURNAL	Submitted (03-APR-2000) to the DDBJ/EMBL/GenBank databases. Yasushi Shintani, Takeda Chemical Industries, LTD, Discovery Research Laboratories 1; 10 Wadai, Tsukuba, Ibaraki 300-4293, Japan (E-mail:Shintani.Yasushi@takeda.co.jp, Tel:81-298-64-5011, Fax:81-298-64-5000)
FEATURES	Location/Qualifiers
source	1..1248
gene	/organism="Homo sapiens" /db_xref="taxon:9606"
CDS	10..1248 /gene="TGR-1" 10..1248 /gene="TGR-1" /codon_start=1 /product="G protein-coupled receptor TGR-1" /protein_id="BAB13721.1" /db_xref="GI:10257381" translation="MEKLNQNSNVYQOKLEDPEOKHLNSTEEYLAFICGRSRHFELPVSVVYPFVVGIVGNVLCLLVILQHOMAKTPNYKFLSVASDLVLLGLMPELVEWMRRYPLEFGVCYFKTALFEYGFASIIITVSERVVALIHPRAKLGSTRARALRIIGWGFSGVLESLPNSTHGIKFPNGSLIPSCATYVKPKMIYNFIIOVTSFLFLYLPMTYSVLYLMALRKDKRLDEAGNANORCRSYNMPLVLVAFLCMAPHIDRLFESFVEEMSESLSAVENTLVHVSSVFYTLSSANNPIYINLSRFQAFOVNISSEHKQWHSQHDPPAPDARNLFTECHVELTEDIGIQPCOSSMHNSHLPALSSSQMSRTNYOSFFHNKT"
BASE COUNT	269 a 393 c 269 g 317 t
ORIGIN	
Query Match	96.1%; Score 1248; DB 9; Length 1248;
Best Local Similarity	100.0%; Pred. No. 8.3e-249;
Matches 1248; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Dn 27	atgtcagggatggaagaacttcagaatctcccgatctaccagcagaactgaagaat 86
Dn 1	atgttcaggagatggaaaaatttcagatatcttcttgatctacacagaaactgaaagt 60
Dn 87	ccattccgaaaacaccttaaacgcacgcgagagatctcggcctctctcgagacctcg 146
Dn 61	ccatttcgaaaacaccttaacacgacgccagagagatctggcccttcttcggcacctcg 120
Dn 147	cgcagcaactctctcccgcgtctctgtttgatctgcacaacsagagctatgaagacgcc 206
Dn 121	cgcacaccacttcttctccgccgtgctgcatgttcgacgacacggcgtatgaagacgcc 180
Dn 207	attggcaatgtctcgtgtgtgcctgttgtatcttcgcacaacsagagctatgaagacgcc 266
Dn 181	attggcaatgtctcgtgtgtgtgcctgtgttcgacgacacggcgtatgaagacgcc 240
Dn 267	aactctaccctcttagcgttggtgtctcttaacctcccgacctgtctcttgaaatgcc 326
Dn 241	aactctaccctcttagcgttggtgtctcttaacctcccgacctgtctcttgaaatgcc 300
Dn 327	ctgagagctatagagatctggcgcaaatcccttctgtgtcggcgccgttgagctgtac 386
Dn 301	ctgagagctctatagagatctggcgcaaatcccttctgtgtcggcgccgttgagctgtac 360
Dn 387	tccaagacgacctctttagaacgtgtgtcttcgacctcaactctcgaaccaacgctc 446
Dn 361	tccaagacgacctctttagaacgtgtgtcttcgacctcaactctcgaaccaacgctc 420

Db	541	ACACGATGCATGCGATTCAGATTCCACTTCTCCCAATGGCTCTGGTCCAGGT	600
Oy	627	tcggccaccctgtaaggltcaatcaagccatgtygaatctacaattlcaatlcgaagltacc	686
Db	601	TCGGCCACCCTGATCGGTATCAAGGCCAATGGATGATCAAAATTTCATCCAGGTCAAC	660
Oy	687	tcctctctatcttaactctctcccaatgaatgtatcaagtgatgagtcctactaactcaatgga	746
Db	661	TCCTTCTATTTACTCTCTCTCCCAATGACTGTATCATGATGCTCTTACTTACCTCATGGCA	720
Oy	747	ctcaagactaaagaagaacaatctctcttgaggagatgaaaggggaatgcaatatlcacaaga	806
Db	721	CTCAGACTTAAGAAAGACAAATCTCTTGAGGCGAGATGAAGGGGAATGCAAATATTCAAGA	780
Oy	807	ccctcgagaaaatcagttcaacaagaatgctgttctgttcttgcttagtgtttgctatctgt	866
Db	781	CCCTCGAGAAATTCATCATCAACAAGATGCTGTGTGTCTTGTGTCTTACTGTGTGTCTACTGT	840
Oy	867	tgaggcccccctcacatctacacgtacacctctctctcaagctcttgaggagtgaggatgacc	926
Db	841	TGGGCCCCCTTGCACATTACCCGACCTCTTTCACGCTTGTGTGAGAGATGGAGTGAATTC	900
Oy	927	ctgagctgcgtgtgtcaaccctgltccaatgtgtgtgtgtcaagtggtcttcttcaactgagctca	986
Db	901	CTGGCTGCTGTGTTCACACCTCGTCCATGTGTGTCAAGTGTCTCTTCTTACTACGTGACCTCA	960
Oy	987	gtctgtcaaccacatatactataaccatctgtctcgcgcgtctccaggcagatctcagaat	1046
Db	961	GCTGTCAACCCCTTATATCTATTAACCTACTGTCTCTCGCCGCTTCCAGGACACATTCCGAAT	1020
Oy	1047	gtgactctctctctccacaacaagtgagacatcccaagatgacccaagtgtgcaaccctgac	1106
Db	1021	GTGATCTCTTCTTCCACAAACAGTGGCACTGCCACATGACACACACAGTTCGCCACTCGCC	1080
Oy	1107	caggcgagaaatcttctctgtgcaagaatgagcaacttgtgtgagctgaacggaaagataagttcc	1166
Db	1081	CAGCGGAACATTTCTCTGTGACAGAAAGCCACTTGTGTGACTGTGACCAAAATATATAGGTGCC	1140
Oy	1167	caattcccaatgtcagttcacatccatgcaactctcaactctcccaacagccctctctagtga	1226
Db	1141	CAATTCCCATGTGACATGCATGCACCAACTGTCACTCCCAACAGCCCTCTCTACTGAA	1200
Oy	1227	cagatgtcagaagaacaactataccaagctctccacttcaacaaaaacttga	1274
Db	1201	CAGATGTCAAGAACAAATCATCAAAACCTTTCACCTTTTAACAAAAACCTGA	1248
RESULT	8		
LOCUS	AF242874		
DEFINITION	Homo sapiens neuromedin U receptor 2 (NMU2R) mRNA, complete cds.		
ACCESSION	AF242874		
VERSION	AF242874.1		
KEYWORDS	human.		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1 (bases 1 to 1239)		
AUTHORS	Howard,A.D., Wang,R., Pong,S.S., Mellin,T.N., Strack,A., Guan,X.M., Zeng,Z., Williams,D.L. Jr., Feigheuer,S.D., Nunes,C.N., Murphy,B., Steir,J.N., Yu,H., Jiang,Q., Clements,M.K., Tan,C.P., McKee,K.K., Hreniuk,D.L., McDonald,T.P., Lynch,K.R., Evans,J.F., Austin,C.P., Caskey,C.T., Van der Ploeg,L.H. and Liu,Q.		
TITLE	Identification of receptors for neuromedin U and its role in feeding		
JOURNAL	Nature 406 (6791), 70-74 (2000)		
MEDLINE	20351041		
REFERENCE	2 (bases 1 to 1239)		
AUTHORS	Liu,Q., McDonald,T.P., Wang,R., Jiang,Q. and Howard,A.D.		
TITLE	Direct Submission		
JOURNAL	Submitted (09-MAR-2000) Pharmacology, Merck Research Labs, West		

FEATURES	POINT, PA 19486, USA
source	Location/Qualifiers 1..1239 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="5"
gene	1..1239 /gene="NM02R"
CDS	1..1239 /gene="NM02R" /note="FM-4" /codon_start=1 /product="neuromedin u receptor 2" /protein_id="AAF02755.1" /db_xref="GI:9082156"
BASE COUNT	267 a 392 c 265 g 315 t
ORIGIN	
Query Match	95.5%; Score 1239; DB 9; Length 1239;
Best Local Similarity	100.0%; Pred. NO. 6.1e-247;
Matches 1239; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
1	ATGGAAGAACTCTCAGAAATGCTCTTGATCTACGACGAGAACTGAAGATCATTTCCAG 60
36	atggaagaaactctcagaatgctcttcgtagatctacacgagaaactagaagaatccattccag 95
1	ATGGAAGAACTCTCAGAAATGCTCTTGATCTACGACGAGAACTGAAGATCATTTCCAG 60
96	aaacacactgaaagacacacgagagagatctgagccttcctctcggaacctcgagcgagcaac 155
61	AAACACCTGGAACAGACGAGAGAGATCTGGGCTTCTCTGGGACCTTGGCGGACGAC 120
156	tctctcccccctgctcgtgtagtgatgtagtgacatgttcttgtagtgtagggtcatctgcaat 215
121	TTCTTCTCTCCCGCTGCTGTGTGTGTGTATGTGCAATTTTGTGTGTGGGTGTCATTGGCAAT 180
216	gtccgtgtagtcctgtagtattctgcagacacgaagctatgaagaagcccaactactaac 275
181	GTCCGTGAGTGTGCTGTGATTTCTGCAGACACGAGCTATGAAGACCCACCAACATCTAC 240
276	ccttcacagcctgagcgctcctcgaactccctggtgcctgcgccttggaatgccccctggaagtc 335
241	CTCTTCACAGCTGGCGGTCTGTGACTCTCTGCTGTGCTCTTGGAAATGCCCTTGAGAGTTC 300
336	tatgagatgtgtagcaactacccttcttgctcggccgttgaggtgtagtctactcaagaagc 395
301	TATGAGATGTGAGCAACATCTCTTGTGTGTGTGGGCGCGGCTGCTACTTCAAGACG 360
396	ggccctcttgaaacggtgctgcgcgcacacccacacacacacacacacacacacacacacacac 455
361	GCCCTCTTGAACCGTGTGCTGCCCTTCACATCTCACACATACCAACCGTACGCTGGAG 420
456	cgctacgtgagcatctcaacacccgtctcgcgcgcacacacacacacacacacacacacacacac 515
421	CGCTACGTGAGCATCTTACACCCGTTCCGGGCGCAACGACAGACACCGGGCGGGGCC 480
516	ctcagagatcctcgagcatcgctcggaggtcttcgcgtgctcttcctccgctcccaacaacagc 575
481	CTCAGAGATCTCTCGCATGTGTGGGGCTTCTCCGCTCTTCTCCCTGCCACACACAGC 540
576	atcagatgcataaattccaaacttcccaatgggttcctcggtccagaggttcagacacacacac 635
541	ATCCATGGCATCAATCTCACTTCCCAATGGGTCCCTGGGTCCAGGTTGGGCCACG 600
636	tgtacggtcatcaagcccaatgtgtagctctacaattctcatcatcaggtacactcctctcta 695

Db 601 TGTACGGTCATCAAGCCATGATGATCTACAAATTCATCATCCAGGTCACTCTTCTTA 660
Qy 696 TCTGCTCTCTCCAGTGTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 755
Db 661 TTTCTACCTCTCTCCAGTGTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
Qy 756 AAGAAAGAAATCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 815
Db 721 AAGAAAGAAATCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
Qy 816 AATCAGTCAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 875
Db 781 AATCAGTCAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
Qy 876 TCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 935
Db 841 TTTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
Qy 936 TGTGTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 995
Db 901 GTGTTCAACCTGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
Qy 996 CCGCAT 1055
Db 961 CCGCAT 1020
Qy 1056 TCTTCCAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1115
Db 1021 TCTTCCAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
Qy 1116 ATCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1175
Db 1081 ATCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
Qy 1176 TGTGTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1235
Db 1141 TGTGTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
Qy 1236 AAGAAAGAAATCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1274
Db 1201 AAGAAAGAAATCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1239

RESULT 9
AX109242 1594 bp DNA PAT 30-APR-2001
LOCUS AX109242 9 from Patent WO0125269.
DEFINITION AX109242
ACCESSION AX109242
VERSION AX109242.1 GI:13924115
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1594)
AUTHORS Delorenz J, W., Berger, C., Loeken, C., Nys, G. and Venema, J.
TITLE Human g-protein coupled receptor
JOURNAL Patent: WO 0125269-A 9 12-APR-2001;
SOLVAY Pharmaceuticals B.V. (NL)
FEATURES
source
1. 1594
/organism="Homo sapiens"
/db_xref="taxon:9606"
55. 945
/note="IGS4A truncated DNA long version"
/codon_start=1
/protein_id="CAC37775.1"
/db_xref="GI:13924116"
/translation="MSGMRKLNAMSYOQKLEDPQKHLNTEYLAFICPRSHF
FLPVSIVYPIFVGVIGNVLCLVILQAKTPPNYLFSLVADLVLKMPLE
VEMRNYPFLPGVGYFKTALFEVCPASILSTVSEYVAVILHFRKLOSTR
RRALRIIGVGFVSLFPLNPSINGIKRHPNPSLVGSAVIAKIPMWTNFIQ
VTSFLYLLPMTVIVSYLYLMLRLKDKSLPADSGNANIQPCRKSVKMLSRSG

BASE COUNT 370 a 473 c 332 g 419 t
ORIGIN
VNPMILCSTSSMNCQVSSST"
Query Match 89.4%; Score 1160; DB 6; Length 1594;
Best Local Similarity 95.1%; Pred. No. 1.4e-230;
Matches 1234; Conservative 0; Mismatches 0; Indels 64; Gaps 1;
Qy 1 AAGGAGAGGCTCAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 60
Db 29 AAGGAGAGGCTCAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 88
Qy 61 GATCAGGAG 120
Db 89 GATCAGGAG 148
Qy 121 ATCTGCT 180
Db 149 ATCTGCT 208
Qy 181 ATGTGCAATTTTGT 240
Db 209 ATGTGCAATTTTGT 268
Qy 241 AGCAGGAG 300
Db 269 AGCAGGAG 328
Qy 301 TCTGTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
Db 329 TCTGTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 388
Qy 361 TCTGTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
Db 389 TCTGTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 448
Qy 421 TCTGTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
Db 449 TCTGTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 508
Qy 481 TCTGTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
Db 509 TCTGTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 568
Qy 541 GCTTCTGCT 600
Db 569 GCTTCTGCT 628
Qy 601 TCCCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
Db 629 TCCCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 688
Qy 661 TCTGTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
Db 689 TCTGTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 748
Qy 721 TCAAGTCT 780
Db 749 TCAAGTCT 808
Qy 781 ATGAGAG 840
Db 809 ATGAGAG 865
Qy 841 TCTGTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
Db 866 TCTGTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 865
Qy 901 GCTTCTGCT 960
Db 866 GCTTCTGCT 924

Query Match	Best Local Similarity	Matches 1344	Conservative	Score 1160	DB 6	Length 1594
Matches 1344	Conservative	0	Mismatches	0	Indels	Gaps
1	aggagagagctgagccttgatcttaatctcagagatggaanaacttcgaatgctcct	60				
29	acggcagagctcagccttgatcttaatctcagagatggaanaacttcgaatgctcct	88				
61	ggatctacacagcagaataatagatccatccatcagaacaaccttgacagcagcagagat	120				
89	ggatctacacagcagaataatagatccatccatcagaacaaccttgacagcagcagagat	148				
121	atctgagcctctctctgaggaactcggcgagcgaactctctcccgctgctgtgtgt	180				
961	cagagctctctctctacagtgagctagctgtgtcaaccatatactactactgtctc	1020				
925	cagagctctctctctacagtgagctagctgtgtcaaccatatactactactgtctc	984				
1021	gacgcttcacagcagcatctccagaatgtgcatctctctccacaacaacagtgacatcc	1080				
985	gacgcttcacagcagcatctccagaatgtgcatctctctccacaacaacagtgacatcc	1044				
1081	agcatgaccacacagctgtgcacactcgcacgcgagacatctctccagaatgtgcatct	1140				
1045	agcatgaccacacagctgtgcacactcgcacgcgagacatctctccagaatgtgcatct	1104				
1141	tgagagctgacagcagaatagatagctcccaatctccatctgtcagatccatctgcaactc	1200				
1105	tgagagctgacagcagaatagatagctcccaatctccatctgtcagatccatctgcaactc	1164				
1201	acctcccaacagcctctctctagtgagacagatgtcaagacaaactcacaacgctccact	1260				
1165	acctcccaacagcctctctctagtgagacagatgtcaagacaaactcacaacgctccact	1224				
1261	ttaacaaaacttgaaatctctctcagaagctgacatctcctc	1298				
1225	tttaacaaaacttgaaatctctctcagaagctgacatctcctc	1262				
RESULT 10	AX109244	1594 bp	DNA	PAT	30-Apr-2001	
LOCUS	AX109244	Sequence 11 from Patent WO0125269.				
DEFINITION	AX109244					
ACCESSION	AX109244.1	GI:13924117				
VERSION						
KEYWORDS	human.					
SOURCE	human.					
ORGANISM	Homo sapiens					
REFERENCE	Human 9-protein coupled receptor					
AUTHORS	Deleersnijder, W., Berger, C., Loeken, C., Nys, G. and Venema, J.					
TITLE	Human 9-protein coupled receptor					
JOURNAL	Patent: WO 0125269-A 11 12-Apr-2001;					
FEATURES	Solway Pharmaceuticals B.V. (NL)					
source	Location/Qualifiers					
	1..1594					
	/organism="Homo sapiens"					
	/db_xref="taxon:9606"					
	64..945					
	/note="IGS4A truncated DNA short version"					
	/codon_start=1					
	/protein_id="CAC37776.1"					
	/db_xref="GI:13924118"					
	/translation="MEKLNASWITVOOKLEDPEQKHLNSTEEYLAFLGPRSHFLPL					
	VSVVYVPIFFVGVIGVNLVCLVLIQHQMKPTNYVLESLAVSDLVLLGLMPLVEVE					
	MMRNPPELFGVGVGFKTALFETVCFASILSTIVSEYVAIILHPRAKQSTRRA					
	LRIGIVMGFVSLPMTSHGIRKPHFPMGSLVPSATCTVIRKPMIYVFIQVS					
	FLFILPMPTVSVLYLYLMALRLKKDKSLSEADBNANIQRPCRSVNMKLSIMBSGVNPR					
	WLLSTSSMKCVSSST"					
BASE COUNT	370 a 473 c 332 g 419 t					
ORIGIN						

Dd			ATCTGGCCCTTCTCTGCGACCTCAGCGCACACTTTCTCTCCCGGTGTCGTGGTG	208
Oy	181	a t g t c c a a t t c t g l b g l y g g l c a t l y g c a t l g c t g l g t c t g l y g a t t c t g c		240
Dd	209	A T G T G C C A A T T T T T G T G T G G G G T C A T T T G C A A T T C T G T G T G C C T G G A T T T G C		268
Oy	241	a g c a c c a g g c t a g a a g a c c c a c a a c t a c t c t c t c a a c c t g y c g l t c t c a a c		300
Dd	269	A G C A C A C G G G C A T G A M A B A C C C C A C C A C T A C T A C T T T C A G C C T G C G G C G T C T T G A C C		328
Oy	301	t c c g y g t c c t g c t c c t t g a a l b c c c t i g a g t l c a l p a g a l y g y g c a a c t a c c t t		360
Dd	329	T C T G G T C T G C T C T T T G G A A T G C C C C T G A G G T C A T A G A T T G G G C A A C T A C C T T		388
Oy	361	t c t t g t l c g y g c c c g l y g s t c t g t a c t l c a a g a c y g c c c t c t l t p a g a c g l t g t c t g		420
Dd	389	T C T T G T T G S G G C C G T G G G T G C T A C T A T A A G A C G S G C C T T T T A G A A C G T G T C T T G C		448
Oy	421	c c t c c a t c c t c a g c a l c a c a c c g l c a e g l y g a g c t a g l y g c a t c c t a c c t c a c c g t		480
Dd	449	C C T C A C T C T C A G A T A C C A C C G T C A G C G T G A G G C G T A C T G G C A T C T C A C C C G T		508
Oy	481	t c c g c g c a a a c t b g c a a g a c c c g y g c c g y g c c c t a g a l t c t t o g a t c g f t c t g		540
Dd	509	T C C G G G C A A A C T b c A A G A G A C C C G G G C C G G G C C T C A G A T C C T C G G C A T C G T C T G G		568
Oy	541	g t c t c t c g l b g t c t l t c c c t g c c c a a c c e a c c a g a l c c a l t g a c l c a a g t l c a c t a c t		600
Dd	569	G C T T C T C G T C T C T T T C C C T G C C C A A C A C A C A G A T C A T G A T G A A T T C A A G T T C A C T A C T		628
Oy	601	t c c c c a a t y g t c c c t g l y g l c c c a g t l c g y c a c t g t a e g t c a t c a a g c c c a l y g a		660
Dd	629	T C C C C A A T G G T C C C T G G T C C A G A T T C G G C C A C C G T A C G G T C A T A A G C C C A A G T G A		688
Oy	661	t c t a a a t t t a c a l c c a a g l t a c c c t c c t a t c t a c t c t c c c a a g a t g t c a		720
Dd	689	T C T A C A A T T T A T A T C A A G T C A C C T C C T T A T T T A C C T C C C C A T G A C T G T C A		748
Oy	721	t c a g t l c t c t a c t a c c l a t g c a c t c a g a c t a a g a a a g a c a a t c t c t t g a g c a g		780
Dd	749	T C A G T G T C T C T A C T A C C T A T G G A C A C T A G A C T A A A A G A A C A A A T C T T T G A G C A G		808
Oy	781	a t g a a g g a a l g c a a t a t c c a a g a c c t g c a g a a a t c a g l c a a c a a g a t c t g t t g		840
Dd	809	A T G A A G G A A A G C A A A A T T T C A A A G A C C C G C A G A A A A C A O T C A A C A A G A T G C Y G T ---		865
Oy	841	t c t t g t c t a g t l g t g c a t c t g l y g g c c c g l t c a c a t l y a c c g a c t c t c t c a		900
Dd	866	-----		865
Oy	901	g c t t g l y g a g a g l y g a g l y a a l c c c t g y g t c l g l t l c a a c c t g l c c a t g t g l g t		960
Dd	866	- C T T T G T G A G A G A T G A G A T C C C T G G C T G T G T T A A A C C T G C A T G H G H G T		924
Oy	961	c a g t g t c t c t c t a c c t a g c t c a g c t g t c a a c c c a t a t c a t a a c t a c t a g t c t c		1020
Dd	925	C A G G T G T C T T T T A C C T A G C T A C C T G T A A C C C A T T A T C A T A A C C A C T A G T G T C		984
Oy	1021	g c c c t t c a g a g a c t t c a a a l y g a t c t c t c t t c c a a a a a g t g y a c t c c c		1080
Dd	985	G C C G C T T T C A G G C A G C A T T C C A G A A T G A T C T T C T T T C A C A A A C A G T G G C A C T C C C		1044
Oy	1081	a g a l t a a c c c a a g t l y g c a c c t l y c c c a g a g a a c t c t c t g a c a g a t l y c c a t t g		1140
Dd	1045	A G C A T A G A C C C A C A T T T G C A C A C T G C C A G G G A N A C T T T C C G A G A G A T G C A C A T T T G		1104
Oy	1141	t g a g a c t y a a c c g a a t a t a g t t c c c a a l t c c a t y t c a g t a l c a l c a l y g c a a c t c t c		1200
Dd	1105	T G S A G T C A C C G A G A T A T A G T G T C C C A A T T C C A A T G T C A G T A C A C A T G C A C A C T C T C		1164
Oy	1201	a c c t c c c a a g c c c t c t a g l y g a a a g a t g l c a a g a a a c t a t a a a g t t c a c t		1260

Db 1165 ACCTCCCAACAGCCCTCTCTAGTGAAGATGTAAGAACAACTATCAAGCTTCCACT 1224

Qy 1261 ttacacaaactgaattcttcagagctgaactcttc 1298
|||||
Db 1225 TTAACAAACCTGAATTCCTTTCAGAGCTGACTCTCTC 1262

RESULT 11
AB041229
LOCUS
DEFINITION
AB041229 1314 bp mRNA ROD 21-SEP-2000
Rattus norvegicus mRNA for G protein-coupled receptor TGR-1,
complete cds.
ACCESSION
AB041229.1 GI:10257382
VERSION
G protein-coupled receptor TGR-1.
KEYWORDS
Rattus norvegicus cDNA to mRNA.
SOURCE
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (stiles)
Hosoya, M.; Moriya, T.; Kawamata, Y.; Ohkubo, S.; Fujii, R.; Matsui, H.,
Shintani, Y., Fukusumi, S., Habata, Y., Hinuma, S., Onda, H.,
Nishimura, O. and Fujino, M.
Identification and Functional Characterization of a Novel Subtype
of Neuregulin Receptor
J. Biol. Chem. 275 (38), 29528-29532 (2000)
10887190
2 (bases 1 to 1314)
Fujii, R., Shintani, Y. and Hinuma, S.
Direct Submission
Submitted (03-APR-2000) to the DDBJ/EMBL/GenBank databases. Shuji
Hinuma, Takeda Chemical Industries, LTD, Discovery Research
Laboratories 1, Madai 10, Tsukuba, Ibaraki 300-4293, Japan
(E-mail: Hinuma-Shuji@takeda.co.jp, Tel: 81-298-64-5035,
Fax: 81-298-64-5000)

FEATURES
source
1..1314
Location/Qualifiers
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
127..1314
/gene="TGR-1"
127..1314
/gene="TGR-1"
/codon_start=1
/product="G protein-coupled receptor TGR-1"
/protein_id="BAB13722.1"
/db_xref="GI:10257383"
/translation="MGKLENAWIMHPMKLYLNSTEEYLAHLGPKRSDLSPLSVAY
ALIFLVGMNLCVIVRHQTLKTPNYLFSLAVSDLVLLGMPLEIYEMNH
PELFGVGCYETALFEIVCFASISLVTVSERVVAIVHPRAKLESTRRALPILS
LWMSVYFSLPNTSINGIKFOHPNGSSVPSASCTYKPKMYNLIQTSLPFI
LPMITLSLYILMKRLKRDSELENNKAVANIHRSKRSVTKMLFVLVRAICWTF
HVDRLFSEFVEWETSLAIVENLIHVGVGFYSSAVNPILYNLSRFRAPRNVS
SPTCKMCHPRHPOPOPPAKKIIFLECHLMELTEDAGPOFPQSSIHNTNLMPACAG
EVP"

BASE COUNT 269 a 397 c 305 g 343 t

ORIGIN

Query Match 61.8%; Score 802.8; DB 10; Length 1314;
Best Local Similarity 78.8%; Pred. No. 1.6e-156;
Matches 976; Conservative 0; Mismatches 247; Indels 15; Gaps 1;

Qy 1 agggagagctcagggcttgatatttaattcagggatggaataacttcgaatgctctc 60
|||||
Db 92 AGGGACACCTGTGGTTGTGATTTTAAGCTCACTAATGGAAACTTGAATAATGCTTCT 151
|||||

Qy 61 ggaactcagcagaagaacatcacaagaacacacagcagaagagat 120
|||||
Db 152 GGATCCAC-----GATCCACTCATGAGTACTTGAAACACACAGAGAGAT 196
|||||

Qy 121 atctggccttctctcaggaactcgcgacgaacattctctcccgctgctgtggtg 180
|||||

Db 197 ACTTGGCCACCTGTGTGGACCAACCGCAGTACACTATCCCTCCGGTGTCTGTGGCCT 256
|||||

Qy 181 agtgcacatttttgggtgggggtcattggcaatgctctgtgtgacctggtatctgc 240
|||||

Db 257 ATGCGCTGATCTTCCGTGGTGGGGTAAATGGCAATCTTGTGGTGCATGATGATTC 316
|||||

Qy 241 agcaccagcctatgaagcagccaccacactactactcttcagcctggggtctgacc 300
|||||

Db 317 GACATGAGACTTTGAAGACACCACCACTACTATCTTTCAGCTTGGAGTCTCAGATC 376
|||||

Qy 301 tccctgctcctcctcttggaatgccccttggaaggtcctatagaatgctggcgaactacc 360
|||||

Db 377 TGCTGTCTCTGCTCTTGGGGATGCGCTGTGAATATCAGAGATGGACAAATTAACCTT 436
|||||

Qy 361 tcttctggcggccgtgggtcgtactactcaagaagcgcctctttagagcgtgtctcg 420
|||||

Db 437 TCCTGTTCGGGCGCTGGGGATGCTACTTCAGACACCCCTCTTCGAGACTGTGCTTTG 496
|||||

Qy 421 cctccactcctcagcactcacaccgctcagctgagagcgtcagctggccactcccaaccgt 480
|||||

Db 497 CCTCCATTCTCAGTGTCCACACGCTTAGCGTAAAGCGCTATGTGGCATTTGCCACCTT 556
|||||

Qy 481 tccgcgcaaacctgcagagcaccgcgcgcgcgcctcagaatcctcgcgacatcgtctgg 540
|||||

Db 557 TCGAGCCCAAGGTGAGAGACGACGCGGCGACGGCCCTCAGATTCCTCAGCTAGTCTGA 616
|||||

Qy 541 gcttctcgctctctctcctcccgccaacacgaatcctcagatcaatgctccactact 600
|||||

Db 617 GCTTCCTGTGTGCTTTTCTTTTGGCCAAATACAGCATCCATGGCATCAAGTTCACAGACT 676
|||||

Qy 601 tcccaatggtccctcgc 660
|||||

Db 677 TTCACAGGCTCTCTCCGATACCTGAGCTCAGCCACGCACTGCACCAACCAACCACTGGG 736
|||||

Qy 661 tctacaattctcactcagcagcagcctcctcactcactcctcccaatgctcgaatgta 720
|||||

Db 737 TGTATTAATCTGATCAATCCAGCTACAGCTTCTCTCTTCAATCCCTCCAAATGACCTCA 796
|||||

Qy 721 tcaagctcctcactcactcactcactcactcactcactcactcactcactcactcactc 780
|||||

Db 797 TCAGCCTCTCTACTACCTCATGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTG 856
|||||

Qy 781 atgaagaggaatgcaaatatcgaagaccctcgaagaataatcgaagaagaatgagcgttg 840
|||||

Db 857 ACAAACTGCTGTGAATATTCACAGACCTCTAGAAAGTCACTCAACCAAGATGCTGTTGG 916
|||||

Qy 841 tcttggtcttaagtgtgtcactcgtgttggtgcccgtctccacaatgacgaactcctca 900
|||||

Db 917 TCTGTGCTCTGCTGTTTGGCATCTGCTGACCCCTTCCATGTGACCGGCTCTCTTCA 976
|||||

Qy 901 gcttctgagagagtgagtggaatccctcgtcgtcgtgtgtcactcgtcgaatgctcgtgtg 960
|||||

Db 977 GCTTTGTGAAAGAGGACAGAGTCCCTGGCTGCTGTGTTCAACCTCATCATGCTGTAT 1036
|||||

Qy 961 caggtgtctctctcactcagcagcagcgtcgtcgaacccatctcctaaactactgctc 1020
|||||

Db 1037 CAGGTCTCTTCTTTATCTGAGCTCGCGCTCAACCCATTAATTAACCTCTGCTCTC 1096
|||||

Qy 1021 gccgcttccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1080
|||||

Db 1097 GCGGCTTCGGGCGGCGCTTTCGAATGTGTCTCCCTACCTGCAAAATGTTGCCATCC 1156
|||||

Qy 1081 agcatgacccaagctgtcgaactcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1140
|||||

Db 1157 GGCATTCAGCAGAGGACCTCCAGCCCAAGATATCTTCTTGAAGATGTCACTCA 1216
|||||

Qy 1141 tgaagctgacgaagatataagttccccaatcctcagatgcaatgcaatcctcagcagc 1200
|||||

Db 1217 TGGAGCTGACAGAGATGAGGCCCCCAATTCCTGCTGATGATTCATTCACACACCA 1276
|||||

Qy 1201 acctcccaacagcctctcagtgaaacagatgccaaga 1238
|||||

Matches 725; Conservative 0; Mismatches 2; Indels 0; Gaps 0;				
Qy	30	tcagggatggaagaaacttcagaaatgcttcctgga	tcacacagagaagaaactagaatcca	89
Db	801	TCAGGGATGGAAGAAACTTCAGAAATGCTTCTGGA	TCACACAGAGAAACTAGAAAGATCCA	742
Qy	90	ttccagaagaacacctgaaacagacacagagat	ctctcctctcctctcctcctcctcctcctcctc	149
Db	741	TTCCAGAGAAACACCTGAACACACGAGAGAT	CTCTCCTCTCCTCCTCCTCCTCCTCCTCCTCCT	682
Qy	150	aggaactctctctcccgctgctgctgctgctg	ctgctgctgctgctgctgctgctgctgctgctg	209
Db	681	AGGCACTCTCTCCCTCCCGGCTGCTGCTGCT	ATGTCATGTCGCAATTTTGTGGGGGCTATT	622
Qy	210	ggcaatgctctgctgctgctgctgctgctgct	gctgctgctgctgctgctgctgctgctgctgct	269
Db	621	GGCAATGCTCTGCTGCTGCTGCTGCTGCTGCT	GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	562
Qy	270	tactaaccttcagcctgctgctgctgctgctg	ctcctcctcctcctcctcctcctcctcctcctc	329
Db	561	TACTACCTTTCAGCCCTGGGGCTCTGACCT	CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	502
Qy	330	gaagctcatagatgctgctgctgctgctgctg	ctgctgctgctgctgctgctgctgctgctgctg	389
Db	501	GAGGTCATGAGATGCTGGCCAACTACCCCT	TTCTGCTGGGGCTGCTGCTGCTGCTGCTGCT	442
Qy	390	aagaagcctctcttgagacgctgctgctgctg	ctcctcctcctcctcctcctcctcctcctcctc	449
Db	441	AAGAAGCCCTCTTTGAGACCGTGTGCTT	GCTTCCCTCCATCCTCAGCATACACACCGT	382
Qy	450	gtgagagcctgctgctgctgctgctgctgctg	ctcctcctcctcctcctcctcctcctcctcctc	509
Db	381	GTGAGAGCGCTACGTCGCAATCTACACCG	CTGCGCGCAAACTGAGAGACACCGCGCG	322
Qy	510	cgggacctgagatccctcctgctgctgctgctg	ctcctcctcctcctcctcctcctcctcctcctc	569
Db	321	CGGGCCCTCAGGATCCCTGGCAATGCTGG	GGCTTCTGCTGCTGCTGCTGCTGCTGCTGCT	262
Qy	570	acagacatcatgcatcaagcttcccaacta	ctcccaatggttccctcctcctcctcctcctc	629
Db	261	ACCAAGATCATGCAATCAAGTTCACACT	TCCCAATGGTCCCTGGTCCCAAGTTCG	202
Qy	630	gccacatgtaagctlcaaaagcccaatg	tgatcacaatttcaatcacaatgtaagctc	689
Db	201	GCCACATGTAAGCTCAAGCCCAATGTGAT	CTACAAATTTCAATCAAGTCAAGTCACTCC	142
Qy	690	ttcctatctaacctcccccacatgactg	tcatcaagtgctcctcctcctcctcctcctc	749
Db	141	TTCTATTTTACCTCTCCCATGACTGTCAT	AGTGTCTCTACTACTCATGCTGACCTC	82
Qy	750	agactaa 756		
Db	81	AGAGTGA 75		
RESULT 15				
AX139107 729 bp DNA PAT 30-MAY-2001				
LOCUS AX139107 Sequence 1 from Patent EP1090990.				
DEFINITION AX139107				
ACCESSION AX139107				
VERSION AX139107.1 GI:14274786				
KEYWORDS				
SOURCE human.				
ORGANISM Homo sapiens				
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE 1 (bases 1 to 729)				
AUTHORS Harland, L. P.				
TITLE Human g-protein-coupled receptor				
JOURNAL Patent: EP 1090990-A 1 11-APR-2001;				
Pfizer Limited (GB) ; PFIZER INC. (US)				
FEATURES				
Location/Qualifiers				
1..729				
source				

BASE COUNT 135 a 245 c 167 g 182 t				
ORIGIN				
/organism="Homo sapiens"				
/db_xref="taxon:9606"				
Query Match 55.3%; Score 717.8; DB 6; Length 729;				
Best Local Similarity 99.7%; Pred. No. 7.3e-139;				
Matches 719; Conservative 0; Mismatches 2; Indels 0; Gaps 0;				
Qy	36	atggaagaaacttcagaaatgcttcctgga	tcacacagagaagaaactagaatcattccag	95
Db	1	ATGGAAGAAACTTCAGAAATGCTTCTGGA	TCACACAGAGAAACTAGAAAGATCCA	60
Qy	96	aaacacctgaaacagacacagagagat	ctcctcctcctcctcctcctcctcctcctcctc	155
Db	61	AAACACCTGAACAGCACCGAGATATG	CTTCTTCTGCGGACTCGGCGGAGCCAC	120
Qy	156	ttcttcctcccgctgctgctgctgctgctg	ctcctcctcctcctcctcctcctcctcctcctc	215
Db	121	TTCTTCTCCCGGCTGCTGCTGCTGCTGCT	GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	180
Qy	216	gtccctgctgctgctgctgctgctgctgctg	ctcctcctcctcctcctcctcctcctcctcctc	275
Db	181	GTCCTGCTGCTGCTGCTGCTGCTGCTGCT	GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	240
Qy	276	ctcttcagcctgagctgctgctgctgctgctg	ctcctcctcctcctcctcctcctcctcctcctc	335
Db	241	CTCTTACGCTTGCGGCTGCTGCTGCTGCT	GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	300
Qy	336	tatgagatgctgctgctgctgctgctgctg	ctcctcctcctcctcctcctcctcctcctcctc	395
Db	301	TATGAGATGCTGCTGCTGCTGCTGCTGCT	GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	360
Qy	396	gccctcttgaagacgctgctgctgctgctg	ctcctcctcctcctcctcctcctcctcctcctc	455
Db	361	GCCCTCTTGTAGACCGCTGCTGCTGCTGCT	CAATCCTCAAGATCAACACCGTACCGTGGAG	420
Qy	456	cgtacgtgacatcctcacaacccgctc	cgcgcgaactgagaagacccgcgcgcgcgc	515
Db	421	CGCTACGTGGCAATCTTACACCGGTTCC	GCGCAAACTGAGAGACACCGCGCGGCGC	480
Qy	516	ctcagatcctcgcgcatcgctgctgctgctg	ctcctcctcctcctcctcctcctcctcctcctc	575
Db	481	CTCAGGATCCTCGGCAATCGTGGGCTT	CCGTGCTTTCCTCGGCCAACACACAG	540
Qy	576	atccatgcatcaagcttcccaact	atccatggttccctcctcctcctcctc	635
Db	541	ATCCATGGGCAATCAAGTTCACCTTCC	CAATGGGCTCCCTGCTCCAGGTTCCGACCC	600
Qy	636	tgtacgtcatcaagcccatgctgcat	ctacatcattcatcatcaccagttcaacctc	695
Db	601	TGTACGTCATCAAGCCCATGTGATCT	ACAAATTTATCATCCAGGTCACTCTTCTTA	660
Qy	696	ttctaacctcctcccaatgcatgct	ctcctcctcctcctcctcctcctcctcctcctc	755
Db	661	TTCTTACCTCTCCCAATGATGATGAT	GATGATGATGATGATGATGATGATGATGATGAT	720
Qy	756	a 756		
Db	721	A 721		

Search completed: April 22, 2002, 18:20:16
Job time: 10975 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 22, 2002, 16:38:20 ; Search time 247.29 Seconds
(without alignments)
4500.013 Million cell updates/sec

Title: US-09-609-146-3

Perfect score: 1298
Sequence: 1 agggagagcctcagccttgg.....tttcagagcctcctcctc 1298

Scoring table:
IDENTITY: NUC
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

N.Geneseq_1101.*
1: /SIDSI/gcgdata/geneseq/geneseqn/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseqn/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseqn/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseqn/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseqn/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseqn/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseqn/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseqn/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseqn/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseqn/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseqn/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseqn/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseqn/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseqn/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseqn/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseqn/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseqn/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseqn/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseqn/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseqn/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseqn/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseqn/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1298	100.0	1298	22	AAD08008
2	1298	100.0	1298	22	AAFB0322
3	1298	100.0	1298	22	AAFB0323
4	1291.6	99.5	1658	22	AAFB0324
5	1291.6	99.5	1658	22	AAFB0325
6	1286.8	99.1	1360	21	AAZ33297
7	1248	96.1	1248	21	AAZ33297
8	1248	96.1	1248	21	AAZ33297
9	1248	96.1	1248	21	AAZ33297
10	1245	95.9	1245	22	AAH45072
11	1238.6	95.4	1245	22	AAH45075

12	1160	89.4	1594	22	AAFB0326
13	1160	89.4	1594	22	AAFB0327
14	799.6	61.6	1231	22	AAD08024
15	787.2	60.6	1185	22	AAH43078
16	723.8	55.8	801	22	AAH50977
17	717.8	55.3	729	22	AAH50977
18	346	26.7	1209	22	AAH45072
19	346	26.7	1209	22	AAH45072
20	346	26.7	1212	21	AAA30663
21	346	26.7	1212	21	AAA30739
22	346	26.7	1318	22	AAD08007
23	344.4	26.5	1212	21	AAZ45707
24	293.6	22.6	1239	22	AAH45076
25	293.6	22.6	1292	22	AAD08025
26	293.6	22.6	1371	22	AAD08026
27	270.6	20.8	1215	22	AAH45073
28	270.6	20.8	1526	21	AAZ45707
29	161.6	12.4	1092	21	AAZ45405
30	161.6	12.4	1092	22	AAFB5450
31	151.2	11.6	1101	21	AAA30732
32	149.6	11.5	1088	18	AAT68664
33	149.6	11.5	1088	18	AAT68756
34	149.6	11.5	1101	21	AAA30643
35	149.6	11.5	1101	21	AAZ51463
36	149.6	11.5	1101	21	AAZ51463
37	144.6	11.1	1063	18	AAT68662
38	144.6	11.1	1063	18	AAT68754
39	142.4	11.0	1050	21	AAZ61492
40	136.4	10.5	1092	18	AAT68760
41	135.8	10.5	1203	22	AAFB5448
42	135	10.4	1098	22	AAFB3681
43	134.8	10.4	1092	22	AAH27800
44	134.8	10.4	1161	22	AAFB3684
45	134.8	10.4	1239	21	AAZ45403

ALIGNMENTS

RESULT	1	
ID	AAD08008	standard; cDNA; 1298 BP.
AC	AAD08008;	
XX		
DT	07-AUG-2001	(first entry)
XX		
DE	Human G-protein coupled receptor, SNORF72 cDNA.	
XX		
KW	Human: G-protein coupled receptor; SNORF72; neuromedin U neuropeptide; NMU; inflammation; arthritis; autoimmune disease; septicemia; psychotic; mental retardation; transplant rejection; neurological disorder; anxiety; respiratory disorder; depression; schizophrenia; dementia; obesity; pain; gastrointestinal disorder; hypertension; epilepsy; diabetes; ischemia; stroke; cancer; sexual disorder; circadian disorder; anorexia; dermatological; psoriasis; Parkinson's disease; nausea; bulimia; allergy; Alzheimer's disease; AIDS; hormonal disorder; memory disorder; migraine; cardiovascular disorder; renal disorder; bone disease; delirium; asthma; Cushing's disease; dysmenorrhea; antiangiinal; cytototoxic; osteoporosis; metabolic disorder; behavioural disorder; Addison's disease; dyskinesia; tranquiliser; antitumor; antiaddictive; ss.	
KW		
OS	Homo sapiens.	
XX		
FT	key	Location/Qualifiers
FT	CDS	27..1274
FT		/*tag= a
FT		/product= "Human SNORF72 receptor"
XX		
XX	WO200144297-A1.	
XX	21-JUN-2001.	
PD		
XX		

Splice variant of
Rat G-protein coup
Nucleotide sequenc
Human nGFR15 codi
Nucleotide sequenc
Human FM-3 coding
Human G-protein co
Human G-protein-co
DNA encoding human
Human G-protein co
Human growth hormo
Human G-protein-co
Human G-protein-co
Human G-protein-co
Rat G-protein coup
Rat G-protein coup
Murine FM-3 coding
Mouse growth hormo
DNA encoding the p
Nucleotide sequenc
DNA encoding human
Human growth hormo
Human growth hormo
Human G-protein-co
Human G-protein-co
Human G-protein-co
Human G-protein-co
Pig growth hormone
Swine growth hormone
cDNA encoding canl
Rat growth hormone
Nucleotide sequenc
Human G-protein co
Rat growth hormone
Short form of mol
cDNA encoding the

PF 13-DEC-2000: 2000MO-US33787.
 XX
 PR 17-DEC-1999: 99US-0466435.
 PR 25-APR-2000: 2000US-0558099.
 PR 30-JUN-2000: 2000US-0609146.
 XX
 PA (SYNA-) SYNAPTIC PHARM CORP.
 PI Boninl JA, Lerman GS, Quan Y, Ogozalek K;
 XX
 DR WPI, 2001-390240/41.
 DR P-PSDB: AAE03629.
 XX
 PT A purified mammalian SNORF62 or SNORF72 receptor protein for
 PT identification of compounds to treat e.g. inflammation, arthritis,
 PT autoimmune diseases, transplant rejection, AIDS, cancer -
 PS
 PS Claim 42: Fig 3; 256pp: English.

XX The invention relates to human G-protein coupled receptors, SNORF62 and
 CC SNORF72 and their corresponding cDNA molecules. SNORF62 and SNORF72
 CC receptors are specific for neuromedin U (NMU) neuropeptides, hence they
 CC are also known as NMU receptors. The agonist and antagonist of NMU
 CC receptors are useful for treating an abnormality in a subject that is
 CC alleviated by decreasing or increasing the activity of NMU receptor.
 CC The NMU receptors serves as a valuable tool for designing drugs which are
 CC useful for treating various pathophysiological conditions such as
 CC inflammation, arthritis, autoimmune diseases, transplant rejection,
 CC graft vs host disease, bacterial, fungal, protozoan and viral infections,
 CC septicemia, AIDS, pain, psychotic and neurological disorders, including
 CC anxiety, depression, schizophrenia, dementia, mental retardation, memory
 CC loss, epilepsy, neuromotor disorders, respiratory disorders, asthma,
 CC eating/body weight disorders including obesity, bulimia, diabetes,
 CC anorexia, nausea, hypertension, hypotension, vascular and cardiovascular
 CC disorders, ischemia, stroke, cancers, sexual disorders, circadian
 CC disorders, renal disorders, bone diseases including osteoporosis, benign
 CC prostatic hypertrophy, gastrointestinal disorders, nasal congestion,
 CC dermatological disorders such as psoriasis, allergies, Parkinson's
 CC disease, Alzheimer's disease, acute heart failure, angina disorders,
 CC delirium and dyskinesias such as Huntington's disease. They can also be
 CC used to regulate steroid hormone disorders, epinephrine release
 CC disorders, electrolyte balance disorders, endocrine disorders, memory
 CC disorders, somatosensory disorders, metabolic disorders, behavioural
 CC disorders, drug addiction, migraine, Addison's disease, Cushing's
 CC disease, prevent miscarriage, induce labour or to treat dysmenorrhoea.
 CC The present cDNA sequence encodes human G-protein coupled receptor,
 CC SNORF72.
 CC
 XX
 SO Sequence 1298 BP: 278 A; 405 C; 282 G; 333 T; 0 other;

Query Match 100.0%; Score 1298; DB 22; Length 1298;
 Best Local Similarity 100.0%; Pred. No. 2, 3e-301;
 Matches 1298; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 agggagagctcagagccttgatcttaatgcagggatggaanaactcagaatgctcct 60
 |||||
 DB 1 agggagagctcagagccttgatcttaatgcagggatggaanaactcagaatgctcct 60
 QY 61 ggattctcagcaggaactggaatcattccagaaacccggaagacccgagagat 120
 |||||
 DB 61 ggattctcagcaggaactggaatcattccagaaacccggaagacccgagagat 120
 QY 121 atctgacctctctcgtcgagacctcgcgagcagcactctctccctgctgctgctg 180
 |||||
 DB 121 atctgacctctctcgtcgagacctcgcgagcagcactctctccctgctgctgctg 180
 QY 181 atctgccaattcttgctgctgctgctgctgctgctgctgctgctgctgctgctg 240
 |||||
 DB 181 atctgccaattcttgctgctgctgctgctgctgctgctgctgctgctgctgctg 240
 QY 241 agcaacgggtatgaaagcccaactactactcctcagcctgctgctgctgctgac 300
 |||||

DB 241 agcaacgggtatgaaagcccaactactactcctcagcctgctgctgctgac 300
 QY 301 tcttggtccctgctccttggaatgccccctggaatctatgagatgctggaactaccc 360
 |||||
 DB 301 tcttggtccctgctccttggaatgccccctggaatctatgagatgctggaactaccc 360
 QY 361 tcttggtccctgctccttggaatgccccctggaatgccccctggaatgctggaact 420
 |||||
 DB 361 tcttggtccctgctccttggaatgccccctggaatgccccctggaatgctggaact 420
 QY 421 cctccatcctcagacatcacaccgctcagcgctgagcgctagctgagcactacaccgc 480
 |||||
 DB 421 cctccatcctcagacatcacaccgctcagcgctgagcgctagctgagcactacaccgc 480
 QY 481 tccggcgcaacacgagagaccccgcgccgccccctcagatctcccgagatgctg 540
 |||||
 DB 481 tccggcgcaacacgagagaccccgcgccgccccctcagatctcccgagatgctg 540
 QY 541 gctctcgtgctctctcctcctgccccacacagacatcacctgagcactacactact 600
 |||||
 DB 541 gctctcgtgctctctcctcctgccccacacagacatcacctgagcactacactact 600
 QY 601 tccccaatggtcccgctgctccagctgctgagcactgagcactacacagcactgag 660
 |||||
 DB 601 tccccaatggtcccgctgctccagctgctgagcactgagcactacacagcactgag 660
 QY 661 tctcaaatctcatcatcacagatctacccctctctatctccctcccccagactgct 720
 |||||
 DB 661 tctcaaatctcatcatcacagatctacccctctctatctccctcccccagactgct 720
 QY 721 tcaagtctcctctactactactcactgagcactcagactaaagaaacaaatctctgag 780
 |||||
 DB 721 tcaagtctcctctactactactcactgagcactcagactaaagaaacaaatctctgag 780
 QY 781 atgaaggaatgaaatataatcaagacccctcgagaatcaatgacacaaagatgct 840
 |||||
 DB 781 atgaaggaatgaaatataatcaagacccctcgagaatcaatgacacaaagatgct 840
 QY 841 tcttggtcctgagctgctgctgctgctgctgctgctgctgctgctgctgctgct 900
 |||||
 DB 841 tcttggtcctgagctgctgctgctgctgctgctgctgctgctgctgctgctgct 900
 QY 901 gcttggtgagagctgagatgaaatccctgctgctgctgctgctgctgctgctgct 960
 |||||
 DB 901 gcttggtgagagctgagatgaaatccctgctgctgctgctgctgctgctgctgct 960
 QY 961 caagtgtcctctctcctcctcctcctcctcctcctcctcctcctcctcctcctc 1020
 |||||
 DB 961 caagtgtcctctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 1020
 QY 1021 gccgcttcagagcagcattccaaatgctgctcctctctcctcctcctcctcctcct 1080
 |||||
 DB 1021 gccgcttcagagcagcattccaaatgctgctcctctcctcctcctcctcctcctc 1080
 QY 1081 agcatgacccagctgagcctcctcctcctcctcctcctcctcctcctcctcct 1140
 |||||
 DB 1081 agcatgacccagctgagcctcctcctcctcctcctcctcctcctcctcctcct 1140
 QY 1141 tggagctgacgagagatagatgctcccaatctcccatgctgctgctgctgctgct 1200
 |||||
 DB 1141 tggagctgacgagagatagatgctcccaatctcccatgctgctgctgctgctgct 1200
 QY 1201 acctcccaaacgacctctcctcctcctcctcctcctcctcctcctcctcctcct 1260
 |||||
 DB 1201 acctcccaaacgacctctcctcctcctcctcctcctcctcctcctcctcctcct 1260
 QY 1261 ttaacaaacacgaaatctctcagagctgagctcctc 1298
 |||||
 DB 1261 ttaacaaacacgaaatctctcagagctgagctcctc 1298

RESULT 2
 AAF80322

ID AAF80322 standard; DNA; 1658 BP.
XX AAF80322;
AC
XX 29-JUN-2001 (first entry)
DT
XX
DE Nucleotide sequence of a G-protein coupled receptor IGS4A long version.
XX
KW Human: G-protein coupled receptor; IGS4; IGS4B; schizophrenia;
KW nervous system disorder; psychiatric disorder; Parkinson's disease;
KW episodic paroxysmal anxiety; phobia; migraine; epilepsy; bulimia; stroke;
KW cardiovascular disease; heart failure; angina pectoris; obesity; emesis;
KW motility disorder; myocardial infarction; hypertension; dyslipidemia;
KW gastrointestinal disorder; inflammatory bowel disease; osteoporosis;
KW inflammation; infection; pain; cancer; immune disorder; allergy; sepsis;
KW gynecological disorder; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 55..1302
FT /tag-
FT /product- "G-protein coupled receptor IGS4A long version"
XX
XX WO200125269-A2.
XX
XX 12-APR-2001.
XX
XX 25-SEP-2000; 2000WO-EP09584.
XX
XX 24-SEP-1999; 99EP-0203140.
XX 24-SEP-1999; 99NL-1013140.
XX 28-JUL-2000; 2000EP-0202683.
XX 31-JUL-2000; 2000US-0222047.
XX
XX (SOLV) SOLVAY PHARM BV.
XX
XX Deleena J de W, Berger C, Loeken C, Nys G, Venema J;
XX
XX WPI: 2001-273568/28.
XX P-PSDB: AAB67802.
XX
XX New G-protein coupled receptors and the polynucleotides encoding them,
XX useful for preventing, ameliorating or correcting nervous system
XX disorders, cardiovascular diseases, dyslipidemias, inflammations, pain
XX or cancers
XX
XX Claim 1: Page 75-77; 102pp; English.
XX
XX The present sequence encodes the long version of a human G-protein
XX coupled receptor designated IGS4A. IGS4 exists in two polymorphic forms,
XX IGS4A and IGS4B. The IGS4 receptors and IGS4 polynucleotides are useful
XX for preventing, ameliorating or correcting dysfunctions or diseases.
XX These diseases include peripheral nervous system, psychiatric and central
XX nervous system disorders (e.g. schizophrenia, episodic paroxysmal
XX anxiety, phobia, Parkinson's disease, migraine, epilepsy, bulimia or
XX stroke), cardiovascular diseases (e.g. heart failure, angina pectoris,
XX myocardial infarction or hypertension), dyslipidemias, obesity, emesis,
XX gastrointestinal disorders (e.g. inflammatory bowel disease or motility
XX disorders), osteoporosis, inflammations, infections (e.g. bacterial,
XX fungal, protozoan or viral), pain, cancers, immune disorders, allergies,
XX sepsis or gynecological disorders. Agonists or antagonists of IGS4 are
XX effective with regard to disorders of the nervous system, including the
XX central and peripheral nervous systems, disorders of the gastrointestinal
XX system, cardiovascular system, skeletal muscle, thyroid, lung or
XX genitourinary system, or immunological disease. The IGS4 polynucleotides
XX are useful as diagnostic reagents for detecting under-expression,
XX overexpression or altered expression of IGS4.
XX
XX Sequence 1658 BP; 377 A; 490 C; 346 G; 445 T; 0 other;

Query Match

100.0%; Score 1298; DB 22; Length 1658;

Best Local Similarity 100.0%; Pred. No. 2,5e-301;
Matches 1298; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 aggggaagctcagcgtctggaatttaattgcaaggatgtaaaaaactcagaatgctctc 60
DB 29 aggggaagctcagcgtctggaatttaattgcaaggatgtaaaaaactcagaatgctctc 88
QY 61 ggaatcaccagcagaactgaagaatccatccagaacaccggaacagcagaagagat 120
DB 89 ggaatcaccagcagaactgaagaatccatccagaacaccggaacagcagaagagat 148
QY 121 atctggtccctcctctgagaccctcgcgccagcactctctccctcctctggtg 180
DB 149 atctggtccctcctctgagaccctcgcgccagcactctctccctcctctggtg 208
QY 181 atgtgccaatcttggtggtggtggtcaattgcaatgctcgtgtgtgctgtgattctgc 240
DB 209 atgtgccaatcttggtggtggtggtggtcaattgcaatgctcgtgtgtgattctgc 268
QY 241 agcacagcgtatgaagagcgcacacacacacacacacacacacacacacacacacac 300
DB 269 agcacagcgtatgaagagcgcacacacacacacacacacacacacacacacacacac 328
QY 301 tctctgtcctcgtcctctggaatgccccctggaaggtctatgaatggtggtggtggt 360
DB 329 tctctgtcctcgtcctctggaatgccccctggaaggtctatgaatggtggtggtggt 388
QY 361 tctctgtcgtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtg 420
DB 389 tctctgtcgtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtg 448
QY 421 cctccac 480
DB 449 cctccac 508
QY 481 tccgagcacaac 540
DB 509 tccgagcacaac 568
QY 541 gctctcgtcgtcctctcctcctcctcctcctcctcctcctcctcctcctcctcctc 600
DB 569 gctctcgtcgtcctctcctcctcctcctcctcctcctcctcctcctcctcctcctc 628
QY 601 tccccaatggtggtcgtggtggtggtggtggtggtggtggtggtggtggtggtggt 660
DB 629 tccccaatggtggtcgtggtggtggtggtggtggtggtggtggtggtggtggtggt 688
QY 661 tctacaattcatcatcaggtcactcctctctatctcctcctcctcctcctcctcctc 720
DB 689 tctacaattcatcatcaggtcactcctctctatctcctcctcctcctcctcctcctc 748
QY 721 tcaagtctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 780
DB 749 tcaagtctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 808
QY 781 atgaagggaatgaatataatcaagaacctcgagaaatcaatgaatgaatgaatgaat 840
DB 809 atgaagggaatgaatataatcaagaacctcgagaaatcaatgaatgaatgaatgaat 868
QY 841 tcttggtcctaggttctgctatctgtggtggtggtggtggtggtggtggtggtggt 900
DB 869 tcttggtcctaggttctgctatctgtggtggtggtggtggtggtggtggtggtggt 928
QY 901 gctttgtggaaggtggaatggaatccctggtggtggtggtggtggtggtggtggtg 960
DB 929 gctttgtggaaggtggaatggaatccctggtggtggtggtggtggtggtggtggtg 988
QY 961 caggtgtcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 1020
DB 989 caggtgtcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 1048
QY 1021 gccggtcagcagcattccagaatgtaattctcttctccacaacagtggtgactcc 1080

Db 1049 gccgcctccagcagcagcattccagatgtgatctctcttccacaaacagtcgactccc 1108
Oy 1081 agcatgaccacagatgcccacccctccagcagcagacatctctcgaagaatgcaacttg 1140
Db 1109 agcatgaccacagatgcccacccctccagcagcagacatctctcgaagaatgcaacttg 1168
Oy 1141 tggagctgaccgagaatatagtgtcccaatcccatgtcagtcacatgcacaaacttc 1200
Db 1169 tggagctgaccgagaatatagtgtcccaatcccatgtcagtcacatgcacaaacttc 1228
Oy 1201 acctcccaacagccctctcagtgaaacagatgtcaagaacaaactatcaagcttccact 1260
Db 1229 acctcccaacagccctctcagtgaaacagatgtcaagaacaaactatcaagcttccact 1288
Oy 1261 ttacaaaacctgaatctcttccagagctgactccctc 1298
Db 1289 ttacaaaacctgaatctcttccagagctgactccctc 1326

RESULT 3
ID AAF80323 standard; DNA; 1658 BP.
AC AAF80323;
XX
XX 29-JUN-2001 (first entry)
DE Nucleotide sequence of G-protein coupled receptor IGS4A short version.
XX
XX Human: G-protein coupled receptor; IGS4; IGS4A; IGS4B; schizophrentia;
XX nervous system disorder; psychiatric disorder; Parkinson's disease;
XX episodic paroxysmal anxiety; phobia; migraine; epilepsy; bulimia; stroke;
XX cardiovascular disease; heart failure; angina pectoris; obesity; emesis;
XX motility disorder; myocardial infarction; hypertension; dyslipidemia;
XX gastrointestinal disorder; inflammatory bowel disease; osteoporosis;
XX inflammation; infection; pain; cancer; immune disorder; allergy; sepsis;
XX gynecological disorder; ss.
OS
XX Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 64..1302
FT /tag= a
FT /product= "G-protein coupled receptor IGS4A short
FT version"
XX
XX W0200125269-A2.
XX
XX 12-APR-2001.
XX
XX 25-SEP-2000; 2000MO-EP09584.
XX
XX 24-SEP-1999; 99EP-0203140.
XX 24-SEP-1999; 99NL-1013140.
XX 28-JUL-2000; 2000EP-0202683.
XX 31-JUL-2000; 2000US-0222047.
XX
XX (SOLV) SOLWAY PHARM BV.
XX
XX Deleersnijder W, Berger C, Loeken C, Nys G, Venema J;
XX
XX WPI: 2001-273568/28.
XX P-PSDB; AAB67803.
XX
XX New G-protein coupled receptors and the polynucleotides encoding them,
XX useful for preventing, ameliorating or correcting nervous system
XX disorders, cardiovascular diseases, dyslipidemias, inflammations, pain
XX or cancers
XX
XX Claim 1: Page 79-81; 102pp; English.
XX
XX The present sequence encodes the short version of a human G-protein
XX coupled receptor designated IGS4A. IGS4 exists in two polymorphic forms,
XX

CC IGS4A and IGS4B. The IGS4 receptors and IGS4 polynucleotides are useful
CC for preventing, ameliorating or correcting dysfunctions or diseases.
CC These diseases include peripheral nervous system, psychiatric and central
CC nervous system disorders (e.g. schizophrenia, episodic paroxysmal
CC anxiety, phobia, Parkinson's disease, migraine, epilepsy, bulimia or
CC stroke), cardiovascular diseases (e.g. heart failure, angina pectoris,
CC myocardial infarction or hypertension), dyslipidemias, obesity, emesis,
CC gastrointestinal disorders (e.g. inflammatory bowel disease or motility
CC disorders), osteoporosis, inflammations, infections (e.g. bacterial,
CC fungal, protozoan or viral), pain, cancers, immune disorders, allergies,
CC sepsis or gynecological disorders. Agonists or antagonists of IGS4 are
CC effective with regard to disorders of the nervous system, including the
CC central and peripheral nervous systems, disorders of the gastrointestinal
CC system, cardiovascular system, skeletal muscle, thyroid, lung or
CC genitourinary system, or immunological disease. The IGS4 polynucleotides
CC are useful as diagnostic reagents for detecting under-expression,
CC overexpression or altered expression of IGS4.
XX
XX Sequence 1658 BP; 377 A; 490 C; 346 G; 445 T; 0 other;

Query Match 100.0%; Score 1298; DB 22; Length 1658;
Best Local Similarity 100.0%; Pred. No. 2,5e-301;
Matches 1298; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 aggggagcctcagccttgatatttaatgcaaggatgaaacattcagaatgcttcc 60
Db 29 aggggagcctcagccttgatatttaatgcaaggatgaaacattcagaatgcttcc 88
Oy 61 ggalctacccagcagaacactagaatccatccagaacacctgaaacagcaggagt 120
Db 89 ggalctacccagcagaacactagaatccatccagaacacctgaaacagcaggagt 148
Oy 121 atcggcctcctctcggagcctcggcagcagcctctctccctgctgtctgtgt 180
Db 149 atcggcctcctcctcggagcctcggcagcagcctctctccctgctgtgtgtgt 208
Oy 181 atgtgccaatcttgt 240
Db 209 atgtgccaatcttgt 268
Oy 241 agcacagcgtatagaagcgcacacacacacacacacacacacacacacacacac 300
Db 269 agcacagcgtatagaagcgcacacacacacacacacacacacacacacacacac 328
Oy 301 tccgtgctcctgctccttggaatgccccttgagagctatgagatgtggcgaacacac 360
Db 329 tccgtgctcctgctccttggaatgccccttgagagctatgagatgtggcgaacacac 388
Oy 361 tctgttcgggcccgt 420
Db 389 tctgttcgggcccgt 448
Oy 421 cctccatcctcagatcac 480
Db 449 cctccatcctcagatcac 508
Oy 481 tccggcgcacaaatgcaagacacacacacacacacacacacacacacacacacacac 540
Db 509 tccggcgcacaaatgcaagacacacacacacacacacacacacacacacacacacac 568
Oy 541 gctctcgcgtgtctctctcctgcccacacacacacacacacacacacacacacacac 600
Db 569 gctctcgcgtgtctctctcctgcccacacacacacacacacacacacacacacacac 628
Oy 601 tcccaaatggttcctgt 660
Db 629 tcccaaatggttcctgt 688
Oy 661 tctacaattcatcatcacaagtcacacacacacacacacacacacacacacacacac 720
Db 689 tctacaattcatcatcacaagtcacacacacacacacacacacacacacacacacac 748


```

Oy 721 tcagtcctctactaccatcagcagctcagactaaagaacaaatctctgagcag 780
    |||
Db 749 tcagtcctctactaccatcagcagctcagactaaagaacaaatctctgagcag 808
Oy 781 atgaaggagatgcaaatatccaagaacccctgcagaaaatcagtaacagatgcttg 840
    |||
Db 809 atgaaggagatgcaaatatccaagaacccctgcagaaaatcagtaacagatgcttg 868
Oy 841 tcttggtcttaagtgttctgtatctgttggcccccgttccacatgacgacttcttca 900
    |||
Db 869 tcttggtcttaagtgttctgtatctgttggcccccgttccacatgacgacttcttca 928
Oy 901 gcttggtgaggagtgagtgaaatccctctgctgctgctgttcaaccctgcacatgctgt 960
    |||
Db 929 gcttggtgaggagtgagtgaaatccctctgctgctgctgttcaaccctgcacatgctgt 988
Oy 961 caggtgctctctctactacgtcagctcagctgtcacaaccatatactataaccctactgtc 1020
    |||
Db 989 caggtgctctctctactacgtcagctcagctgtcacaaccatatactataaccctactgtc 1048
Oy 1021 gccgcctcagagcagatccccaagatgatactctcttcccaaacagtgacactcc 1080
    |||
Db 1049 gccgcctcagagcagatccccaagatgatactctcttcccaaacagtgacactcc 1108
Oy 1081 agcatgaccacagatgacacccctgcagcagagacatcttccctgacagatgacacttg 1140
    |||
Db 1109 agcatgaccacagatgacacccctgcagcagagacatcttccctgacagatgacacttg 1168
Oy 1141 tggagtcgacagagatagatagctcccaatcccaatgacatgacatgacacacttc 1200
    |||
Db 1169 tggagtcgacagagatagatagctcccaatcccaatgacatgacatgacacacttc 1228
Oy 1201 accctcccaacagcgcctctgagtgacacagatgacagaaacacttcaaaagcttccact 1260
    |||
Db 1229 accctcccaacagcgcctctgagtgacacagatgacagaaacacttcaaaagcttccact 1288
Oy 1261 ttaacaaaacccgtaattcttccagagctgactctc 1298
    |||
Db 1289 ttaacaaaacccgtaattcttccagagctgactctc 1326

RESULT 4
AAF80324
ID AAF80324 standard; DNA; 1658 BP.
XX
AC AAF80324;
XX
DT 29-JUN-2001 (first entry)
XX
DE Nucleotide sequence of a G-protein coupled receptor IGS4B long version.
XX
KW Human; G-protein coupled receptor; IGS4; IGS4A; IGS4B; schizophrenia;
    nervous system disorder; psychiatric disorder; Parkinson's disease;
    episodic paroxysmal anxiety; phobia; migraine; epilepsy; bulimia; stroke;
    cardiovascular disease; heart failure; angina pectoris; obesity; emesis;
    motility disorder; myocardial infarction; hypertension; dyslipidemia;
    gastrointestinal disorder; inflammatory bowel disease; osteoporosis;
    inflammation; infection; pain; cancer; immune disorder; allergy; sepsis;
    gynecological disorder; ss.
KW Homo sapiens.
XX
OS Homo sapiens.
XX
FI Key Location/Qualifiers
FT CDS 55..1302
    /tag= a
    /product= "G-protein coupled receptor IGS4B long version"
XX
PN WO200125269-A2.
XX
PD 12-APR-2001.
XX
PF 25-SEP-2000; 2000WO-EP09584.
XX

```

```

PR 24-SEP-1999; 99EP-0203140.
PR 24-SEP-1999; 99NL-1013140.
PR 28-JUL-2000; 2000EP-0202683.
PR 31-JUL-2000; 2000US-0222047.
XX
PA (SOLV ) SOLVAY PHARM BV.
XX
PI Deleersnijder W, Berger C, Loeken C, Nys G, Venema J:
DR WPI: 2001-273568/28.
DR P-PSDB: AAB67804.
XX
PT New G-protein coupled receptors and the polynucleotides encoding them,
PT useful for preventing, ameliorating or correcting nervous system
PT disorders, cardiovascular diseases, dyslipidemias, inflammations, pain
PT or cancers
XX
PS Claim 1; Page 83-85; 102pp; English.
XX
CC The present sequence encodes the long version of a human G-protein
CC coupled receptor designated IGS4B. IGS4 exists in two polymorphic forms,
CC IGS4A and IGS4B. The IGS4 receptors and IGS4 polynucleotides are useful
CC for preventing, ameliorating or correcting dysfunction or diseases.
CC These diseases include peripheral nervous system, psychiatric and central
CC nervous system disorders (e.g. schizophrenia, episodic paroxysmal
CC anxiety, phobia, Parkinson's disease, migraine, epilepsy, bulimia or
CC stroke), cardiovascular diseases (e.g. heart failure, angina pectoris,
CC myocardial infarction or hypertension), dyslipidemias, obesity, emesis,
CC gastrointestinal disorders (e.g. inflammatory bowel disease or motility
CC disorders), osteoporosis, inflammations, infections (e.g. bacterial,
CC fungal, protozoan or viral), pain, cancers, immune disorders, allergies,
CC sepsis or gynecological disorders. Agonists or antagonists of IGS4 are
CC effective with regard to disorders of the nervous system, including the
CC central and peripheral nervous systems, disorders of the gastrointestinal
CC system, cardiovascular system, skeletal muscle, thyroid, lung or
CC genitourinary system, or immunological disease. The IGS4 polynucleotides
CC are useful as diagnostic reagents for detecting under-expression,
CC overexpression or altered expression of IGS4.
XX
SQ Sequence 1658 BP; 378 A; 489 C; 346 G; 445 T; 0 other:

Query Match 99.5%; Score 1291.6; DB 22; Length 1658;
Best Local Similarity 99.7%; Pred. No. 8.4e-300;
Matches 1294; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy 1 agggagagctcagccttgatttaatgctcagagatggaacaaactcagaatgctcct 60
    |||
Db 29 agggagagctcagccttgatttaatgctcagagatggaacaaactcagaatgctcct 88
Oy 61 ggaatcaccagcagaactagaatgccatccagaacaccccgaaacagcagcaggaggt 120
    |||
Db 89 ggaatcaccagcagaactagaatgccatccagaacaccccgaaacagcagcaggaggt 148
Oy 121 atctggccttcctctcgagaccccgagcagcaactctcccccgtctgtgtgtg 180
    |||
Db 149 atctggccttcctctcgagaccccgagcagcaactctcccccgtctgtgtgtg 208
Oy 181 atgtgccaatcttctgtgtgtgtgtcatgtgcaatgctcctgtgtgtgtgtc 240
    |||
Db 209 atgtgccaatcttctgtgtgtgtgtcatgtgcaatgctcctgtgtgtgtgtc 268
Oy 241 agcaccaagctatgaagagcccaacaactacactctcagcctggcggtctcgacc 300
    |||
Db 269 agcaccaagctatgaagagcccaacaactacactctcagcctggcggtctcgacc 328
Oy 301 tcttggtcctgctccttggaatgcccttggaaggttcatatgagatggttggcaactaccc 360
    |||
Db 329 tcttggtcctgctccttggaatgcccttggaaggttcatatgagatggttggcaactaccc 388
Oy 361 tcttggtcggcccgctggagctgacttcaagacggccctcttgagaccgtgtgtctg 420
    |||
Db 389 tcttggtcggcccgctggagctgacttcaagacggccctcttgagaccgtgtgtctg 448

```

```

OY 421 cctccatcctcagacatcacccgctcagctgagcgctacgtggccatctaacccgt 480
    |||||||
DB 449 cctccatcctcagacatcacccgctcagctgagcgctacgtggccatctaacccgt 508
OY 481 tccgagcacaactgacagagaccccgccgcccctcaagatcctcgcagctcgtg 540
    |||||||
DB 509 tccgagcacaactgacagagaccccgccgcccctcaagatcctcgcagctcgtg 568
OY 541 gctctcgcgctcctcctcctccctccacacccgacatccatctgcatcaagtctccact 600
    |||||||
DB 569 gctctcgcgctcctcctcctccctccacacccgacatccatctgcatcaagtctccact 628
OY 601 tcccaaatggtccctcgttcccgagttccgacacgtttagagttcatcaagcccatgga 660
    |||||||
DB 629 tcccaaatggtccctcgttcccgagttccgacacgtttagagttcatcaagcccatgga 688
OY 661 tctacaattctcatcatccaggtcactccttccatctcactcctccatgactgtca 720
    |||||||
DB 689 tctacaattctcatcatccaggtcactccttccatctcactcctccatgactgtca 748
OY 721 tccggttccctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 780
    |||||||
DB 749 tccggttccctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 808
OY 781 atgaagggaaatgcacaatatccaagaacctgcagagaaatcagttcaacaagatcgtgtt 840
    |||||||
DB 809 atgaagggaaatgcacaatatccaagaacctgcagagaaatcagttcaacaagatcgtgtt 868
OY 841 tcttggtcttagtgttctgtatctgtgtggcccggttccacatctgacagctcttctca 900
    |||||||
DB 869 tcttggtcttagtgttctgtatctgtgtggcccggttccacatctgacagctcttctca 928
OY 901 gcttcttgagaggtgagaggaatccctcgtcgtgtgttcaacctcgttccatggtgtgt 960
    |||||||
DB 929 gcttcttgagaggtgagaggaatccctcgtcgtgtgttcaacctcgttccatggtgtgt 988
OY 961 caggtgtctctcctcactcactgagctcagctgttcaaacccatctatctataactcgtctc 1020
    |||||||
DB 989 caggtgtctctcctcactcactgagctcagctgttcaaacccatctatctataactcgtctc 1048
OY 1021 gccgcttcacagagcagatccacgaatgtgatctctcttccacaaacagttgcaactcc 1080
    |||||||
DB 1049 gccgcttcacagagcagatccacgaatgtgatctctcttccacaaacagttgcaactcc 1108
OY 1081 agcatgaccacacagttgcacccctgcagcagaggaacatctcctgcagagaaatgcaacttt 1140
    |||||||
DB 1109 agcatgaccacacagttgcacccctgcagcagaggaacatctcctgcagagaaatgcaacttt 1168
OY 1141 tggagctgacagagatatagtgtcccaattcccatgtcagttcatctcagttcacaactctc 1200
    |||||||
DB 1169 tggagctgacagagatatagtgtcccaattcccatgtcagttcatctcagttcacaactctc 1228
OY 1201 acctcccaaacagccctctcagtgacacagatgtcacaagaacatcaaatcactccact 1260
    |||||||
DB 1229 acctcccaaacagccctctcagtgacacagatgtcacaagaacatcaaatcactccact 1288
OY 1261 ttaacaaaacctgaattcttctcagaggtgactctcctc 1298
    |||||||
DB 1289 ttaacaaaacctgaattcttctcagaggtgactctcctc 1326

```

RESULT 5

AAF80325
ID AAF80325 standard; DNA; 1658 BP.

XX AAF80325;

XX 29-JUN-2001 (First entry)

XX Nucleotide sequence of G-protein coupled receptor IGS4B short version.

KX Human; G-protein coupled receptor; IGS4; IGS4A; IGS4B; schizophrenia1;

```

KW nervous system disorder; psychiatric disorder; Parkinson's disease;
KW episodic paroxysmal anxiety; phobia; migraine; epilepsy; bulimia; stroke;
KW cardiovascular disease; heart failure; angina pectoris; obesity; emesis;
KW motility disorder; myocardial infarction; hypertension; dyslipidemia;
KW gastrointestinal disorder; inflammatory bowel disease; osteoporosis;
KW inflammation; infection; pain; cancer; immune disorder; allergy; sepsis;
KW gynecological disorder; ss.
XX
OS Homo sapiens.
XX
PH Key Location/Qualifiers
FT CDS 64..1302
FT /tag="a
FT /product="G-protein coupled receptor IGS4B short
FT version"
XX
PN WO200125269-A2.
XX
PD 12-APR-2001.
XX
PF 25-SEP-2000; 2000MO-EP09584.
XX
PR 24-SEP-1999; 99EP-0203140.
PR 24-SEP-1999; 99NL-1013140.
PR 28-JUL-2000; 2000EP-0202683.
PR 31-JUL-2000; 2000US-0222047.
XX
PA (SOLV ) SOLVAY PHARM BV.
XX
PI Deleersnijder W, Berger C, Loeken C, Nys G, Venema J;
XX
XX WPI; 2001-273568/28.
XX P-PSDB; AAB67805.
XX
PT New G-protein coupled receptors and the polynucleotides encoding them,
PT useful for preventing, ameliorating or correcting nervous system
PT disorders, cardiovascular diseases, dyslipidemias, inflammations, pain
PT or cancers -
XX
PS Claim 1; Page 87-89; 102pp; English.
XX
CC The present sequence encodes the short version of a human G-protein
CC coupled receptor designated IGS4B. IGS4 exists in two polymorphic forms,
CC IGS4A and IGS4B. The IGS4 receptors and IGS4 polynucleotides are useful
CC for preventing, ameliorating or correcting dysfunctions or diseases.
CC These diseases include peripheral nervous system, psychiatric and central
CC nervous system disorders (e.g. schizophrenia), episodic paroxysmal
CC anxiety, phobia, Parkinson's disease, migraine, epilepsy, bulimia or
CC stroke), cardiovascular diseases (e.g. heart failure, angina pectoris,
CC myocardial infarction or hypertension), dyslipidemias, obesity, emesis,
CC gastrointestinal disorders (e.g. inflammatory bowel disease or motility
CC disorders), osteoporosis, inflammations, infections (e.g. bacterial,
CC fungal, protozoan or viral), pain, cancers, immune disorders, allergies,
CC sepsis or gynecological disorders. Agonists or antagonists of IGS4 are
CC effective with regard to disorders of the nervous system, including the
CC central and peripheral nervous systems, disorders of the gastrointestinal
CC system, cardiovascular system, skeletal muscle, thyroid, lung or
CC genitourinary system, or immunological disease. The IGS4 polynucleotides
CC are useful as diagnostic reagents for detecting under-expression,
CC overexpression or altered expression of IGS4.
XX
SQ Sequence 1658 BP; 378 A; 489 C; 346 G; 445 T; 0 other;

```

Query Match 99.5%; Score 1291.6; DB 22; Length 1658;
Best Local Similarity 99.7%; Pred. No. 8..4e-300;
Matches 1294; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

OY 1 aggggaagcctcagccttgatcttaatgtcagagatggaacacatcagaatgcttcc 60
    |||||||
DB 29 aggggaagcctcagccttgatcttaatgtcagagatggaacacatcagaatgcttcc 88
OY 61 ggaatctacagcagaactagaagatccatccagaacacatcgaacacgacccagagagt 120

```

```

Db      89 ggaattacacagagaacatagatgccattccagaacaacctgacacagcagagat 148
Oy      121 atcttgccctctctctgagacactcggcgagacactctctctcccgctgtgtgtgt 180
Db      149 atcttgccctctctctgagacactcggcgagacactctctctcccgctgtgtgtgt 208
Oy      181 atgtgccaaattctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 240
Db      209 atgtgtccaaattctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 268
Oy      241 agcacacagagctatgaagacgcccacaaactactacctcttcagcctgtgcgtcttgacc 300
Db      269 agcacacagagctatgaagacgcccacaaactactacctcttcagcctgtgcgtcttgacc 328
Oy      301 tcttggtctctgtctctcttggaatgccccctgtgaggtctctatagatgtgtgtgtgtgt 360
Db      329 tcttggtctctgtctctcttggaatgccccctgtgaggtctctatagatgtgtgtgtgt 388
Oy      361 tcttggtctctgtctctcttggaatgccccctgtgaggtctctatagatgtgtgtgtgt 420
Db      389 tcttggtctctgtctctcttggaatgccccctgtgaggtctctatagatgtgtgtgtgt 448
Oy      421 cctccatctctcagcatcaccaacgctcagcggtgtgagcgctacgltgtgtgtgtgtgt 480
Db      449 cctccatctctcagcatcaccaacgctcagcggtgtgagcgctacgltgtgtgtgtgt 508
Oy      481 tccggtcccaaatctgacagagacacggcgcgccgctcagagatcctctcggcatcgtgtgt 540
Db      509 tccggtcccaaatctgacagagacacggcgcgccgctcagagatcctctcggcatcgtgt 568
Oy      541 gcttctcgtgtctctctctcgtcccaacacacagcatcctatgcatcaggtctccactact 600
Db      569 gcttctcgtgtctctctctcgtcccaacacacagcatcctatgcatcaggtctccactact 628
Oy      601 tccccaatgtgtctcctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 660
Db      629 tccccaatgtgtctcctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 688
Oy      661 tctaaatctctctctctcaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 720
Db      689 tctaaatctctctctcaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 748
Oy      721 tcaagtgtctctctctctcaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 780
Db      749 tcaagtgtctctctctctcaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 808
Oy      781 atgaagggaaatgcaaatatctcaaaagacccctgacagaaatcgaatcgaatgtgtgtgt 840
Db      809 atgaagggaaatgcaaatatctcaaaagacccctgacagaaatcgaatcgaatgtgtgtgt 868
Oy      841 tcttggtctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 900
Db      869 tcttggtctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 928
Oy      901 gcttctgagagaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 960
Db      929 gcttctgagagaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 988
Oy      961 caggtgtctctctctctcaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1020
Db      989 caggtgtctctctctcaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1048
Oy      1021 gctggtctcagagcagatctcagaaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1080
Db      1049 gctggtctcagagcagatctcagaaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1108
Oy      1081 agcatgacacacaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1140
Db      1109 agcatgacacacaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1168
Oy      1141 tggagctgacacgaagatataaggttccccaattcccatgtcagatcatccatgacaaactctc 1200

```

```

Db      1169 tggagctgacacgaagatataaggttccccaattcctatgtatcagtcacgtgtgacaaactctc 1228
Oy      1201 acctcccaacacacccctctctatgtgaacagatgtcaagaacaaactataaagcttccact 1260
Db      1229 acctcccaacacacccctctctatgtgaacagatgtcaagaacaaactataaagcttccact 1288
Oy      1261 ttaacaaacctgaattctcttcagagctgacatctctc 1298
Db      1289 ttaacaaacctgaattctcttcagagctgacatctctc 1326

RESULT 6
AA233297
ID AA233297 standard; cDNA; 1360 BP.
XX
AC AA233297;
XX
DT 21-FEB-2000 (first entry)
XX
DE Human neurotensin-like receptor encoding cDNA.
XX
KW Human; neurotensin-like receptor; NLR: G-protein coupled receptor;
KW central nervous system; anesthesia; analgesia; ds.
XX
OS Homo sapiens.
XX
PN W09955732-A1.
XX
PD 04-NOV-1999.
XX
PE 15-APR-1999; 99WO-SE00598.
XX
PR 24-APR-1998; 98SE-0001455.
XX
PA (ASTR-) ASTRA PHARMA INC.
PA (ASTR ) ASTRA AB.
XX
PI Ahmad S, Cao J, O'Donnell D, Walker P;
XX
DR WPI: 2000-052803/04.
DR P-PSDB: AAY52992.
XX
PT Novel neurotensin-like receptor, useful for identifying agents for
PT producing anesthesia or analgesia
XX
PS Claim 11; Fig 1; 46pp; English.
XX
CC The present sequence encodes a human G-protein coupled receptor
CC neurotensin-like receptor (NLR). The NLR polynucleotide and protein
CC can be used to isolate compounds that bind, (ant)agonise or alter
CC the activity or expression of the NLR. The NLR is a G-protein coupled
CC receptor which is expressed in the central nervous system and shares
CC homology with human neurotensin receptor. The receptors can be used in
CC assays to identify agents for producing anesthesia and analgesia.
XX
SO Sequence 1360 BP; 293 A; 423 C; 294 G; 350 T; 0 other;

Query Match          99.1%; Score 1286.8; DB 21; Length 1360;
Best Local Similarity 99.5%; Pred. No. 1.le-298;
Matches 1291; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Oy      1 agggagagctcagagctgtgatatttaatgtcagggatgtgaaaaacttcagaatgtctct 60
Db      39 agggagagctcagagctgtgatatttaatgtcagggatgtgaaaaacttcagaatgtctct 98
Oy      61 ggatcaccagcagaacacagagaagatccattccagaanaacctgtgacacagcaggagat 120
Db      99 ggaattacacagagaacacagagaagatccattccagaanaacctgtgacacagcaggagat 158
Oy      121 atctgctctctctgtcggaacctcggcgacgaactctctctcccggtgtgtgtgtgt 180
Db      159 atctgctctctctgtcggaacctcggcgacgaactctctctcccggtgtgtgtgtgt 218

```

[illegible]

Oy	1261	ttacaacacctgaattcttccagactgactctc	1298
Dd	1299	ttacaacacctgaattcttccagactgactctc	1336
RESULT# 7			
AD01123	ID	AD01123 standard; cDNA; 1248 BP.	
XX	AC	AD01123;	
XX	DT	02-NOV-2000 (first entry)	
XX	DE	Human orphan G protein-coupled receptor hrup6 cDNA.	
XX	KW	Human; orphan G protein-coupled receptor; GPCR; hrup6; drug screening;	
XX	KX	transmembrane receptor; signal cascade; ss.	
OS	Homo sapiens.		
XX	Key	Location/Qualifiers	
FH	FT	1..1248	
FT	CDS	/tag= a	
FT		/product= "hrup6"	
FT		/note= "human orphan G protein-coupled receptor"	
Pn		WO200031258-A2.	
XX		02-JUN-2000.	
PD		13-OCT-1999; 99WO-US23687.	
PF			
XX		20-NOV-1998; 98US-0109213.	
PR		16-FEB-1999; 99US-0120416.	
PR		26-FEB-1999; 99US-0121852.	
PR		12-MAR-1999; 99US-0123946.	
PR		12-MAR-1999; 99US-0123949.	
PR		28-MAY-1999; 99US-0136436.	
PR		28-MAY-1999; 99US-0136437.	
PR		28-MAY-1999; 99US-0136543.	
PR		28-MAY-1999; 99US-0136567.	
PR		28-MAY-1999; 99US-0137127.	
PR		28-MAY-1999; 99US-0137131.	
PR		29-JUN-1999; 99US-0141448.	
PR		29-SEP-1999; 99US-0156555.	
PR		29-SEP-1999; 99US-0156653.	
PR		29-SEP-1999; 99US-0156654.	
PR		29-SEP-1999; 99US-0156657.	
PR		01-OCT-1999; 99US-0157280.	
PR		01-OCT-1999; 99US-0157281.	
PR		01-OCT-1999; 99US-0157282.	
PR		01-OCT-1999; 99US-0157293.	
PR		01-OCT-1999; 99US-0157294.	
PR		12-OCT-1999; 99US-0416760.	
PR		12-OCT-1999; 99US-0417044.	
XX			
PA	(AREN-)	ARENA PHARM INC.	
XX			
PI	Chen R,	Dang HT, Liaw CW, Lin I,	
XX			
DR	WPI:	2000-400068/34.	
DR	P-PDB:	AAY71296.	
PT			
XX			
XX			
PS	Claim 21;	Page 56-57; 102pp; English.	
CC			
CC	The present sequence is a cDNA encoding hrup6, an endogenous human		
CC	orphan G protein-coupled receptor (GPCR). The full length hrup6 cDNA		
CC	was cloned by RT-PCR using human thymus cDNA as template.		
CC	The orphan GPCR of the invention, like all GPCRs has seven transmembrane		
CC	alpha helices with an extracellular N-terminus and an intracellular		

CC C-terminus. However, no endogenous ligands has yet been identified for
CC the proteins of the invention. The orphan GPCRs may be used in the
CC identification of their endogenous ligands, and to screen potential GPCR
CC agonists and antagonists for use as pharmaceutical agents. The proteins
CC may also be used in the study of GPCR-mediated signalling cascades, and
CC to elucidate their precise role in normal and diseased human conditions.
CC Nucleic acid encoding human orphan GPCRs may be used for tissue
CC localisation expression analysis to provide information about their
CC function in healthy and pathological states.
CC
xx

SQ Sequence 1248 BP; 269 A; 393 C; 269 G; 317 T; 0 other:

Query Match 96.1%; Score 1248; DB 21; Length 1248;
Best Local Similarity 100.0%; Pred. No. 2,1e-289;
Matches 1248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Oy 27 atgtcaggatggaagaaactctcagaatgctctctggtatctaccagagaactagaat 86
Db 1 atgtcaggatggaagaaactctcagaatgctctctggtatctaccagagaactagaat 60
Oy 87 ccattccagaagaaacacttgaaacagacagagagatctctctctcgtcggaactcgg 146
Db 61 ccattccagaagaaacacttgaaacagacagagagatctctctctcgtcggaactcgg 120
Oy 147 cgcagcgaactctctctcccgctgctgtgtgtatgtatgtgccaattttgtgtgggggtc 206
Db 121 cgcagcgaactctctctcccgctgctgtgtgtatgtatgtgccaattttgtgtgggggtc 180
Oy 207 attggaacatgctcctgtgtcctgtgtatctcgaacagacagcctaagaagcaccac 266
Db 181 attggaacatgctcctgtgtcctgtgtatctcgaacagacagcctaagaagcaccac 240
Oy 267 aactactaactcttcaagcctgtgcgtctctgaactcctctgctcctcttggaatgcc 326
Db 241 aactactaactcttcaagcctgtgcgtctctgaactcctctgctcctcttggaatgcc 300
Oy 327 cctggaagctatagatgctgtgcgaactacccctctctgttctggggccgttggtctgac 386
Db 301 cctggaagctatagatgctgtgcgaactacccctctctgttctggggccgttggtctgac 360
Oy 387 ttcagaagcggccctcttctgaacacgtgtgtctcgcctccatccatccagca tccaccacgtc 446
Db 361 ttcagaagcggccctcttctgaacacgtgtgtctcgcctccatccatccagca tccaccacgtc 420
Oy 447 agcgtggaagcgtatcgtgtgcatactctacacccgcttcgcgcgaactgcaagacacgg 506
Db 421 agcgtggaagcgtatcgtgtgcatactctacacccgcttcgcgcgaactgcaagacacgg 480
Oy 507 cgcgggcccctcaggaatcctcgtcgtctgtggtcttcgcgtgtcttctccctgtccc 566
Db 481 cgcgggcccctcaggaatcctcgtcgtctgtggtcttcgcgtgtcttctccctgtccc 540
Oy 567 aacaccagacatccatgagca tcaagttccactactctcccaatgggtccctgtgcccaggt 626
Db 541 aacaccagacatccatgagca tcaagttccactactctcccaatgggtccctgtgcccaggt 600
Oy 627 tctggccacactgtgaagctcacaagcccatgtgtatctacaatttcaatccaggtacc 686
Db 601 tctggccacactgtgaagctcacaagcccatgtgtatctacaatttcaatccaggtacc 660
Oy 687 tccctctcattctacactctcccca tgaatgcatcagtgctctctcactaactcattagga 746
Db 661 tccctctcattctacactctcccca tgaatgcatcagtgctctctcactaactcattagga 720
Oy 747 ctcaagaactaagaagaagaactctctcttgaggcagatgaaaggagatgcaata tccaaga 806
Db 721 ctcaagaactaagaagaagaactctctcttgaggcagatgaaaggagatgcaata tccaaga 780
Oy 807 cctctgagaagaaatcagtaacagaatgtctgtctgtctgtctgtctgtctgtctgtctgt 866
Db 781 cctctgagaagaaatcagtaacagaatgtctgtctgtctgtctgtctgtctgtctgtctgt 840
```

```
Oy 867 tgggccctctccacatctgaccgaactctctcagctctgtgtgaggagtgtgaatcc 926
Db 841 tgggccctctccacatctgaccgaactctctcagctctgtgtgaggagtgtgaatcc 900
Oy 927 ctgtgcctgtgtgttcaactctgtccatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 986
Db 901 ctgtgcctgtgtgttcaactctgtccatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 960
Oy 987 gctgttaaccccatctatctataactactgtctgtcgcctctcagggcagatctcagaat 1046
Db 961 gctgttaaccccatctatctataactactgtctgtcgcctctcagggcagatctcagaat 1020
Oy 1047 gtgactctctctctccacaacagtggaactctccagcagatgaccagatgtccactgtgc 1106
Db 1021 gtgactctctctctccacaacagtggaactctccagcagatgaccagatgtccactgtgc 1080
Oy 1107 cagcggaacatctctctcgaagaatgccactctgtgtgagctgacccaagatataagttccc 1166
Db 1081 cagcggaacatctctctcgaagaatgccactctgtgtgagctgacccaagatataagttccc 1140
Oy 1167 caattccatgtcatcatcatcatcatgtgacaactctacctccacaagcctctctagtgaa 1226
Db 1141 caattccatgtcatcatcatcatcatgtgacaactctacctccacaagcctctctagtgaa 1200
Oy 1227 cagatgtcaagaacaacatcatcaagcttccacttcaacaaacctga 1274
Db 1201 cagatgtcaagaacaacatcatcaagcttccacttcaacaaacctga 1248

RESULT 8
ID AAA46022
AAA46022 standard; cDNA; 1248 BP.
XX
AC AAA46022:
XX
DT 22-AUG-2000 (first entry)
XX
DE Human G protein coupled receptor hRup6 encoding cDNA SEQ ID NO:11.
KW Human; G protein coupled receptor; GPCR; transmembrane receptor;
KW Identification; agonist; screening; therapeutic; pharmaceutical;
XX mutant; ss.
OS Homo sapiens.
XX
PN W0200022131-A2.
XX
PD 20-APR-2000.
XX
PF 13-OCT-1999; 99MO-US24065.
XX
PR 13-OCT-1998; 98US-0170496.
XX
PR 12-NOV-1998; 98US-0108029.
XX
PR 20-NOV-1998; 98US-0109213.
XX
PR 27-NOV-1998; 98US-0110060.
XX
PR 16-FEB-1999; 99US-0120416.
XX
PR 26-FEB-1999; 99US-0121852.
XX
PR 12-MAR-1999; 99US-0123944.
XX
PR 12-MAR-1999; 99US-0123945.
XX
PR 12-MAR-1999; 99US-0123946.
XX
PR 12-MAR-1999; 99US-0123948.
XX
PR 12-MAR-1999; 99US-0123951.
XX
PR 12-MAR-1999; 99US-0123951.
XX
PR 28-MAY-1999; 99US-0136436.
XX
PR 28-MAY-1999; 99US-0136437.
XX
PR 28-MAY-1999; 99US-0136439.
XX
PR 28-MAY-1999; 99US-0137127.
XX
PR 28-MAY-1999; 99US-0137131.
XX
PR 28-MAY-1999; 99US-0137567.
XX
PR 30-JUN-1999; 99US-0141448.
XX
PR 27-AUG-1999; 99US-0151114.
XX
PR 03-SEP-1999; 99US-0152524.
XX
PR 29-SEP-1999; 99US-0156633.
```

PR 29-SEP-1999; 9905-0156555.
 PR 29-SEP-1999; 9905-0156634.
 XX
 PA (AREN-) ARENA PHARM INC.
 XX
 PI Behan DP, Lehmann-Brulinsma K, Chalmers DT, Chen R, Dang HT;
 PI Gore M, Liaw CW, Lin I, Lowitz K, White C;
 XX
 DR MPI: 2000-317986/27.
 DR P-PSDB: AAB02830.
 XX
 PT Non-endogenous, human G protein-coupled receptors for screening
 PT receptor, inverse or partial agonists useful as therapeutic agents
 PS
 PS Example 1; Page 85-86; 187pp; English.
 CC The present invention describes transmembrane receptors, preferably
 CC human G protein coupled receptors (GPCR), for which the endogenous
 CC ligand is unknown (orphan GPCR receptors). More specifically the present
 CC invention relates to non-endogenous, constitutively activated versions
 CC of a human GPCR. These non-endogenous human GPCRs can be useful for
 CC the direct identification of candidate compounds as receptor agonists,
 CC inverse agonists or partial agonists for use as pharmaceutical agents.
 CC AAA6017 to AAA46126 and AAB02825 to AAB02859 represent sequences used in
 CC the exemplification of the present invention.
 XX
 XX Sequence 1248 BP; 269 A; 393 C; 269 G; 317 T; 0 other;
 SQ
 Query Match 96.1%; Score 1248; DB 21; Length 1248;
 Best Local Similarity 100.0%; Pred. No. 2,1e-289;
 Matches 1248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 27 atgtcaggaatggaagaaactctcagaatgctccttgatctacacagaaactagaat 86
 Db 1 atgtcaggaatggaagaaactctcagaatgctccttgatctacacagaaactagaat 60
 QY 87 ccaatccagaagaacacctgaaacagacagagagatctgagctctctctcgtgagacccg 146
 Db 61 ccaatccagaagaacacctgaaacagacagagagatctgagctctctctcgtgagacccg 120
 QY 147 cgaagacacactctctctcccggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 206
 Db 121 cgaagacacactctctctcccggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 180
 QY 207 atgtgcaatgtctgt 266
 Db 181 atgtgcaatgtctgt 240
 QY 267 aactactaactcttcagcctgtgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 326
 Db 241 aactactaactcttcagcctgtgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 300
 QY 327 ctggagagctctatgagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 386
 Db 301 ctggagagctctatgagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 360
 QY 387 tcaaaagagccctctcttgagacgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 446
 Db 361 tcaaaagagccctctcttgagacgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 420
 QY 447 agcgtgagagcgtacgt 506
 Db 421 agcgtgagagcgtacgt 480
 QY 507 cgcgcggccctcagagatctctcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 566
 Db 481 cgcgcggccctcagagatctctcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 540
 QY 567 aacacacagatcatcatgcatcaagttccaaactctccccaatgggtcccggtccaggt 626
 Db 541 aacacacagatcatcatgcatcaagttccaaactctccccaatgggtcccggtccaggt 600

QY 627 tcggcacctgtacgtcatcaagcccatgttgatctacatctcatccacagttcac 686
 Db 601 tcggcacctgtacgtcatcaagcccatgttgatctacatctcatccacagttcac 660
 QY 687 tccctctatcttaacctctcccaatgactgtcatcagttctccttactactcaatg 746
 Db 661 tccctctatcttaacctctcccaatgactgtcatcagttctccttactactcaatg 720
 QY 747 ctgaactaaagaagaagaatctcttgagacagatgtgaaaggaatgaaatattcaaga 806
 Db 721 ctgaactaaagaagaagaatctcttgagacagatgtgaaaggaatgaaatattcaaga 780
 QY 807 cccctgcagaaatcagtcacaaagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 866
 Db 781 cccctgcagaaatcagtcacaaagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 840
 QY 867 tgggcccgttccacatctgacacgtctcttcaagcttvtgagagagtgagtgaaatcc 926
 Db 841 tgggcccgttccacatctgacacgtctcttcaagcttvtgagagagtgagtgaaatcc 900
 QY 927 ctgagctgtgttcaacctcgtccatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 986
 Db 901 ctgagctgtgttcaacctcgtccatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 960
 QY 987 gctgtcaaccctatctatcaactactgtctcgcgcgttccagagcagatccagaat 1046
 Db 961 gctgtcaaccctatctatcaactactgtctcgcgcgttccagagcagatccagaat 1020
 QY 1047 gtatctctctcttccacaacagtgagactcccaagacatgacacagttgacactgtcc 1106
 Db 1021 gtatctctctcttccacaacagtgagactcccaagacatgacacagttgacactgtcc 1080
 QY 1107 cagcaggaacatctctcagagaagaatgacacttvtgagagtgagccgaagatagatgcc 1166
 Db 1081 cagcaggaacatctctcagagaagaatgacacttvtgagagtgagccgaagatagatgcc 1140
 QY 1167 caattccatgtcagtcacatccacacacactcaccctcccaacagccctctagtgaa 1226
 Db 1141 caattccatgtcagtcacatccacacacactcaccctcccaacagccctctagtgaa 1200
 QY 1227 caagatgtcaagaacaaactatcaaaagcttccactttaaacaacactgaa 1274
 Db 1201 caagatgtcaagaacaaactatcaaaagcttccactttaaacaacactgaa 1248
 RESULT 9
 AAH49526
 ID AAH49526 standard; DNA; 1248 BP.
 XX
 AC AAH49526;
 XX
 DT 21-SEP-2001 (first entry)
 XX
 DE Human GTP-binding protein-coupled receptor GPRV39 coding sequence.
 XX
 KW GTP-binding protein-coupled receptor; neuroprotective; immunomodulatory;
 KW muscular; urinary; circulatory; anorectic; gene therapy; human;
 KW guanosine triphosphate; G-protein; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200148189-A1.
 XX
 PD 05-JUL-2001.
 XX
 PF 28-DEC-2000; 2000WO-JP09409.
 XX
 PR 28-DEC-1999; 99JP-0375152.
 PR 31-MAR-2000; 2000JP-0101339.
 PR 23-MAY-2000; 2000JP-0155978.
 XX
 PA (HELI-) HELIX RES INST.
 XX

PI Matsumoto S, Oda T, Saito Y, Morikawa N, Yoshida K, Suwa M;
 PI Sugiyama T;
 XX
 DR WPI: 2001-425663/45.
 XX P-PSDB: MAG64297.
 XX
 PT Family of guanosine triphosphate binding protein coupled receptors and
 PT genes encoding them for treatment and prevention of diseases associated
 PT with these receptors -
 XX
 PS Claim 1: Page 105: 137pp; Japanese.
 XX
 CC The present sequence is the coding sequence for a human guanosine
 CC triphosphate (GTP)-binding protein-coupled receptor. The receptor is
 CC useful for the investigation, diagnosis, treatment and prevention of
 CC diseases associated with GTP-binding protein-coupled receptors, including
 CC neurological, circulatory, digestive system, immune system, muscle and
 CC urinary system disorders. GTP-binding proteins are also known as
 CC G-proteins.
 XX
 SQ Sequence 1248 BP: 269 A; 393 C; 269 G; 317 T; 0 other;

Query Match 96.1%; Score 1248; DB 22; Length 1248;
 Best Local Similarity 100.0%; Pred. No. 2,1e-289;
 Matches 1248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 27 atgtgaaggatgaaaactcagaatgttcttgatctaccagagaactagaagt 86
DB 1 atgtcaggatgtgaaaactcagaatgttcttgatctaccagagaactagaagt 60
OY 87 ccattccagaagaacactgagacagacccagagatctgacctctctcgagacctcg 146
DB 61 ccattccagaagaacactgagacagacccagagatctgacctctctcgagacctcg 120
OY 147 cgcgagccattctctcccgctgctgtgtgtatgtgccaatttctgtgtgtgtgtc 206
DB 121 cgcgagccattctctcccgctgctgtgtgtgtatgtgccaatttctgtgtgtgtc 180
OY 207 attggcaatgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 266
DB 181 attggcaatgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 240
OY 267 aactactacactctcagacctgagctgtctctcagacctctcagacctctcagacct 326
DB 241 aactactacactctcagacctgagctgtctctcagacctctcagacctctcagacct 300
OY 327 ctggaagctctatgagatgtgtgagcaactaccttctctgtctgtgtgtgtgtgtgt 386
DB 301 ctggaagctctatgagatgtgtgagcaactaccttctctgtctgtgtgtgtgtgtgt 360
OY 387 ttcaagagacgacctcttttgaagacgctgtgtgtgtgtgtgtgtgtgtgtgtgtgt 446
DB 361 ttcaagagacgacctctcttgaagacgctgtgtgtgtgtgtgtgtgtgtgtgtgtgt 420
OY 447 agcgtgtgagcgtactgtgtgacatctctacacccgcttcgcgcacaaatgcagagacc 506
DB 421 agcgtgtgagcgtactgtgtgacatctctacacccgcttcgcgcacaaatgcagagacc 480
OY 507 cgcgagacctcagaatctctcggcatcgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgt 566
DB 481 cgcgagacctcagaatctctcggcatcgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgt 540
OY 567 aacacacagacatctgagcactcaagtctcctacacacacacacacacacacacacac 626
DB 541 aacacacagacatctgagcactcaagtctcctacacacacacacacacacacacacac 600
OY 627 tcggcaccactgtacgtactcaagacacatgtgtgactataatctaccatccacagctacc 686
DB 601 tcggcaccactgtacgtactcaagacacatgtgtgactataatctaccatccacagctacc 660
OY 687 tctctctactatctacacctctccacatgactgcatcaagtgtctctactactatgaca 746
DB 667 tctctctactatctacacctctccacatgactgcatcaagtgtctctactactatgaca 720

```

```

DB 661 tctctctactatctacacctctccacatgactgcatcaagtgtctctactactatgaca 720
OY 747 ctcaagactaaagaagacaaatctcttgaggagatgaaaggaaatgcaaatctcaaga 806
DB 721 ctcaagactaaagaagacaaatctcttgaggagatgaaaggaaatgcaaatctcaaga 780
OY 807 ccttgcaaaaatcagtaacaagaatgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 866
DB 781 ccttgcaaaaatcagtaacaagaatgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 840
OY 867 tgggcccccttcacatctgacacctctctcagacctctgtgtgtgtgtgtgtgtgtgtgt 926
DB 841 tgggcccccttcacatctgacacctctctcagacctctgtgtgtgtgtgtgtgtgtgtgt 900
OY 927 ctggcctgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 986
DB 901 ctggcctgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 960
OY 987 gctgtcaacccatctatcctacacctgtctcgcgcgtctccagcagacattccagaa 1046
DB 961 gctgtcaacccatctatcctacacctgtctcgcgcgtctccagcagacattccagaa 1020
OY 1047 gtgactctctcttccacaacaagatgagcactccacagatgacacacagttgcaacctg 1106
DB 1021 gtgactctctcttccacaacaagatgagcactccacagatgacacacagttgcaacctg 1080
OY 1107 cagcggaacatctctctgacagaatgtccacctgtgtgtgtgtgtgtgtgtgtgtgtgt 1166
DB 1081 cagcggaacatctctctgacagaatgtccacctgtgtgtgtgtgtgtgtgtgtgtgtgt 1140
OY 1167 caattccatctgacatgacacatgacacacacacacacacacacacacacacacacac 1226
DB 1141 caattccatctgacatgacacatgacacacacacacacacacacacacacacacacac 1200
OY 1227 cagatgtcaagaacaactatcaaaagtctcacttaacaaaactgtga 1274
DB 1201 cagatgtcaagaacaactatcaaaagtctcacttaacaaaactgtga 1248

```

```

RESULT 10
AAH43072
ID AAH43072 standard; DNA; 1245 BP.
XX
AC AAH43072;
XX
XX
DT 15-OCT-2001 (first entry)
XX
DE Nucleotide sequence of a human TGR-1 protein.
XX
KW TGR-1; neuromedin U; hypertension; stress disease; ss.
XX
OS Homo sapiens.
XX
XX
FH Key Location/Qualifiers
FT CDS 1..1245
FT /tag= a
FT /product= "TGR-1"
FT /note= "no termination codon given"
XX
XX
PN WO200157524-A1.
XX
PD 09-AUG-2001.
XX
XX
PF 02-FEB-2001; 2001WO-JP00746.
XX
XX
PR 04-FEB-2000; 2000JP-0032773.
PR 24-FEB-2000; 2000JP-005252.
PR 30-MAR-2000; 2000JP-0097896.
PR 19-JUN-2000; 2000JP-0187536.
XX
XX (TAKE ) TAKEDA CHEM IND LTD.
XX
XX Hinuma S, Shintani Y, Hosoya M, Fujii R, Moriya T, Matsui H;
PI

```


PI Okubo S:
 XX WPI: 2001-488917/53.
 DR P-PSDB: AAG63353.
 XX
 PT Identifying predicted or actual structures of two or more members of a
 PT chemical or physical library by mass spectrometry compising
 PT correlating molecular mass measurements of two or more members with a
 PT shared chemical history -
 XX
 PS Claim 9; Page 79-80; 95pp: Japanese.
 XX
 CC The present sequence encodes a human TGR-1 protein. The specification
 CC describes a method of screening a compound, which is capable of binding
 CC properties of neuromedin U to TGR-1. The method is useful for screening
 CC preventatives and remedies for hypertension, stress diseases, etc..
 CC TGR-1 antagonists are also useful for treating the same diseases.
 CC
 XX
 SQ Sequence 1245 BP; 268 A; 393 C; 268 G; 316 T; 0 other;

Query Match 95.9%; Score 1245; DB 22; Length 1245;
 Best Local Similarity 100.0%; Pred. No. 1,1e-288;
 Matches 1245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 atgtcaggagtggaataactcagaatgcttcttgagatctaccagcagaaactagaagat 85
 Db 1 atgtcaggagtggaataactcagaatgcttcttgagatctaccagcagaaactagaagat 60
 QY 87 ccattccagaagaacctgaacaagacacagagagataltctgaccttctctctcggaacctcg 146
 Db 61 ccattccagaagaacctgaacaagacacagagagataltctgaccttctctctcggaacctcg 120
 QY 147 cgcagcacacttctctctcccggtgtctgtgtatgtatgtccaattttctgtgtgtgtgtc 206
 Db 121 cgcagcacacttctctctcccggtgtctgtgtatgtatgtccaatlttctgtgtgtgtgtc 180
 QY 207 attgcgaatgtctctgtgtgtcctgtgtatcttcgaacacagcgtatgaagacgccacc 266
 Db 181 attgcgaatgtctctgtgtgtcctgtgtatcttcgaacacagcgtatgaagacgccacc 240
 QY 267 aactaactaactcttcagccttgccgtgtctctgaacctctgtctctcttggaatgcc 326
 Db 241 aactaactaactcttcagccttgccgtgtctctgaacctctgtctctcttggaatgcc 300
 QY 327 ctggaaggtcctatgagatgtgtgcgaactaactcttcttctgtgtgtgtgtgtgtgtgtac 386
 Db 301 ctggaaggtcctatgagatgtgtgcgaactaactcttcttctgtgtgtgtgtgtgtgtgtac 360
 QY 387 ttcaagacgagcctctcttgagaccgtgtgtctgcgctccatctccacagacacacgctc 446
 Db 361 ttcaagacgagcctctcttgagaccgtgtgtctgcgctccatctccacagacacacgctc 420
 QY 447 agcgtggaagcgtacgtgtgcacactacacccgcttcgcgcgaactgtcagaagaccccg 506
 Db 421 agcgtggaagcgtacgtgtgcacactacacccgcttcgcgcgaactgtcagaagaccccg 480
 QY 507 cgcgcgagcctctcagatccctcgcgcatctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtcc 566
 Db 481 cgcgcgagcctctcagatccctcgcgcatctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtcc 540
 QY 567 aacacacgacatcgaatgacaaagtccacactctcccaatgggtccctgtccacgct 626
 Db 541 aacacacgacatcgaatgacaaagtccacactctcccaatgggtccctgtccacgct 600
 QY 627 tcggcacaactgtacgtgtacaaagcccatgttgatctacaatttcacatccacagtgacc 686
 Db 601 tcggcacaactgtacgtgtacaaagcccatgttgatctacaatttcacatccacagtgacc 660
 QY 687 tccctctatcttaccctccctccacatgacgtatcatatgttctctctaccctcctatggca 746
 Db 661 tccctctatcttaccctccctccacatgacgtatcatatgttctctctaccctcctatggca 720

QY 747 ctcaagaataaagaagacaaatctcttgagcagatgaaggaatgcaaatatcaagaaga 806
 Db 721 ctcaagaataaagaagacaaatctcttgagcagatgaaggaatgcaaatatcaagaaga 780
 QY 807 cccctcagaataatcagtcacaagaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 866
 Db 781 cccctcagaataatcagtcacaagaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 840
 QY 867 tgggccccgttcacattgacacgtctcttcaagcttttgagagatgagatgacac 926
 Db 841 tgggccccgttcacattgacacgtctcttcaagcttttgagagatgagatgacac 900
 QY 927 ctggtctgtgtgttcacacctgtccatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 986
 Db 901 ctggtctgtgtgttcacacctgtccatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 960
 QY 987 gctgtcaaacccatctatctataactactgtctcgcgcgttcaggaagcatccagaat 1046
 Db 961 gctgtcaaacccatctatctataactactgtctcgcgcgttcaggaagcatccagaat 1020
 QY 1047 gtgatctcttcttcacaaacagtggaactcccaagcatgacccacagttgccaactgac 1106
 Db 1021 gtgatctcttcttcacaaacagtggaactcccaagcatgacccacagttgccaactgac 1080
 QY 1107 cagcgaacaactctctcagacagatgccacttctgtgtgtgtgtgtgtgtgtgtgtgtgt 1166
 Db 1081 cagcgaacaactctctcagacagatgccacttctgtgtgtgtgtgtgtgtgtgtgtgtgt 1140
 QY 1167 caattccatgtgtatcatcagtcacaactctcaactcccaacagccctctatgtgaa 1226
 Db 1141 caattccatgtgtatcatcagtcacaactctcaactcccaacagccctctatgtgaa 1200
 QY 1227 caagtgtcaagaacaactatcaaggttccactttcaacaagaacc 1271
 Db 1201 caagtgtcaagaacaactatcaaggttccactttcaacaagaacc 1245

RESULT 11

AAH43075
 ID AAH43075 standard; DNM: 1245 BP.

AC AAH43075;

DT 15-OCT-2001 (first entry)

DE Nucleotide sequence of a human TGR-1 protein.

KW TGR-1; neuromedin U; hypertension; stress disease; ss.

OS Homo sapiens.

FT Key Location/Qualifiers

FT CDS 1..1245 /tag= a

FT /product= "TGR-1"

FT /note= "no termination codon given"

PN MO200157524-A1.

PD 09-AUG-2001.

PF 02-FEB-2001; 2001WO-JP00746.

PR 04-FEB-2000; 2000JP-0032773.

PR 24-FEB-2000; 2000JP-0052252.

PR 30-MAR-2000; 2000JP-0097896.

PR 19-JUN-2000; 2000JP-0187536.

PA (TAKE) TAKEDA CHEM IND LTD.

PI Hinuma S, Shintani Y, Hosoya M, Fujii R, Moriya T, Matsui H;

PI Okubo S;

KW gastrointestinal disorder; inflammatory bowel disease; osteoporosis;
KM inflammation; infection; pain; cancer; immune disorder; allergy; sepsis;
XX gynecological disorder; ss.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 64..945
FT /tag= a
FT /product= "splice variant of G-protein coupled receptor
FT IGS4A short version"
XX
XX WO200125269-A2.
XX
XX 12-APR-2001.
XX
XX 25-SEP-2000: 2000WO-EP09584.
XX
XX 24-SEP-1999: 99EP-0203140.
XX 24-SEP-1999: 99NL-1013140.
XX 28-JUL-2000: 2000EP-0202683.
XX 31-JUL-2000: 2000US-0222047.
XX
XX (SOLV) SOLWAY PHARM BV.
XX
XX Deleernijder W, Berger C, Loeken C, Nys G, Venema J;
XX WPI: 2001-273568/28.
XX P-PSDB: AAB67807.
XX
XX New G-protein coupled receptors and the polynucleotides encoding them,
PT disorders, cardiovascular diseases, dyslipidemias, inflammations, pain
PT or cancers
XX
XX Example 1b: Page 94-95: 102pp: English.
XX
XX The present sequence encodes a splice variant of the short version of a
CC human G-protein coupled receptor designated IGS4A. IGS4 exists in two
CC polymorphic forms, IGS4A and IGS4B. The IGS4 receptors and IGS4
CC polynucleotides are useful for preventing, ameliorating or correcting
CC dysfunctions or diseases. These diseases include peripheral nervous
CC system, psychiatric and central nervous system disorders
CC (e.g. schizophrenia, episodic paroxysmal anxiety, phobia, Parkinson's
CC disease, migraine, epilepsy, bulimia or stroke), cardiovascular diseases
CC (e.g. heart failure, angina pectoris, myocardial infarction or
CC hypertension), dyslipidemias, obesity, emesis, gastrointestinal
CC disorders (e.g. inflammatory bowel disease or motility disorders),
CC osteoporosis, inflammations, infections (e.g. bacterial, fungal,
CC protozoan or viral), pain, cancers, immune disorders, allergies,
CC sepsis or gynecological disorders. Agonists or antagonists of IGS4 are
CC effective with regard to disorders of the nervous system, including the
CC central and peripheral nervous systems, disorders of the gastrointestinal
CC system, cardiovascular disease, skeletal muscle, thyroid, lung or
CC genitourinary system, or immunological disease. The IGS4 polynucleotides
CC are useful as diagnostic reagents for detecting under-expression,
CC overexpression or altered expression of IGS4.
XX
XX
SQ Sequence 1594 BP; 370 A; 473 C; 332 G; 419 T; 0 other;

Query Match 89.4%; Score 1160; DB 22; Length 1594;
Best Local Similarity 95.1%; Pred. No. 2, 7e-268;
Matches 1234; Conservative 0; Mismatches 0; Indels 64; Gaps 1;

QY 1 agggagagcctcagccttgatttaattcagggatgtaaaaaactcagaatgctcct 60
DB 29 agggagagcctcagccttgatttaattcagggatgtaaaaaactcagaatgctcct 88
QY 61 aggtctaccagcagaactgaagatcattccagaacacccctgaaagacccgagagat 120
DB 89 ggtactaccagcagaactgaagatcattccagaacacacccctgaaagacccgagagat 148

QY 121 atctgacctctcctcgcgcagcactcggcgagcaccactctctccctcgctcgtggtg 180
DB 149 atctgacctctcctcgcgcagcactcggcgagcaccactctctccctcgctcgtggtg 208
QY 181 atgtgcgaattttgt 240
DB 209 atgtgcgaattttgt 268
QY 241 agcacagcgtatgaagagcccccacacactactactctctcaacccctgtggtgtgacc 300
DB 269 agcacagcgtatgaagagcccccacacactactactctctcaacccctgtggtgtgacc 328
QY 301 tccctgttccctcctcttgaaatgcccccttgagggttctaagatgtgtgcgaactaccct 360
DB 329 tccctgttccctcctcttgaaatgcccccttgagggttctaagatgtgtgcgaactaccct 388
QY 361 tctgtgtggtggtggt 420
DB 389 tctgtgtggtggtggt 448
QY 421 cctccatccctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 480
DB 449 cctccatccctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 508
QY 481 tccgcgccaacatgcagagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 540
DB 509 tccgcgccaacatgcagagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 568
QY 541 gcttctcgtgtct 600
DB 569 gcttctcgtgtct 628
QY 601 tccccaatgt 660
DB 629 tccccaatgt 688
QY 661 tctacaattcattcattcattcattcattcattcattcattcattcattcattcattcatt 720
DB 689 tctacaattcattcattcattcattcattcattcattcattcattcattcattcattcatt 748
QY 721 tcaagtgtct 780
DB 749 tcaagtgtct 808
QY 781 atgaaggaatgtgaatatttcaagaagccctgcaggaataatgaagcaagaatgtgtgt 840
DB 809 atgaaggaatgtgaatatttcaagaagccctgcaggaataatgaagcaagaatgtgtgt 865
QY 841 tcttgytcttagt 900
DB 866 ----- 865
QY 901 gcttgtggaagagtgagtgaaatccctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 960
DB 866 -cttgtggaagagtgagtgaaatccctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 924
QY 961 caagtgtct 1020
DB 925 caagtgtct 984
QY 1021 gccgcttcagagcagcattccagaatgtgattctctctctctccacaacacagtgagcactcc 1080
DB 985 gccgcttcagagcagcattccagaatgtgattctctctctctccacaacacagtgagcactcc 1044
QY 1081 agcatgaccacagctgtgcaactgtgccagcggaacatctctctgcaagaatgtccactttg 1140
DB 1045 agcatgaccacagctgtgcaactgtgccagcggaacatctctctgcaagaatgtccactttg 1104
QY 1141 tggagctgaccaggaatataatgtgtccccaattctccatgtcaatgtatcatcatgacaaacttc 1200
DB 1105 tggagctgaccaggaatataatgtgtccccaattctccatgtcaatgtatcatcatgacaaacttc 1164
QY 1201 acctcccaagacccctctctagtgaaacagatgtcagaagaacacataaccaagcttccact 1260

PE 13-DEC-2000; 2000WO-US33787.
 XX
 PR 17-DEC-1999; 99US-0466435.
 PR 25-APR-2000; 2000US-0558099.
 PR 30-JUN-2000; 2000US-0609146.
 XX
 PA (SYNA-) SYNAPTIC PHARM CORP.
 XX
 XX
 PI Bonini JA, Lerman GS, Quan Y, Ogozalek K;
 XX
 DR WPI; 2001-390240/41.
 XX P-PSDB; AAE03634.
 DR
 XX
 PT A purified mammalian SNORF62 or SNORF72 receptor protein for
 PT identification of compounds to treat e.g. inflammation, arthritis,
 PT autoimmune diseases, transplant rejection, AIDS, cancer -
 XX
 PS Claim 15; Fig 14; 256pp: English.

The invention relates to human G-protein coupled receptors, SNORF62 and SNORF72 and their corresponding cDNA molecules. SNORF62 and SNORF72 receptors are specific for neuromedin U (NMU) neuropeptides, hence they are also known as NMU receptors. The agonist and antagonist of NMU receptors are useful for treating an abnormality in a subject that is alleviated by decreasing or increasing the activity of NMU receptor. The NMU receptors serves as a valuable tool for designing drugs which are useful for treating various pathophysiological conditions such as inflammation, arthritis, autoimmune diseases, transplant rejection, graft vs host disease, bacterial, fungal, protozoan and viral infections, septicemia, AIDS, pain, psychotic and neurological disorders, including anxiety, depression, schizophrenia, dementia, mental retardation, memory loss, epilepsy, neuromotor disorders, respiratory disorders, asthma, eating/body weight disorders including obesity, bulimia, diabetes, anorexia, nausea, hypertension, hypotension, vascular and cardiovascular disorders, ischaemia, stroke, cancers, sexual disorders, circadian disorders, renal disorders, bone diseases including osteoporosis, benign prostatic hypertrophy, gastrointestinal disorders, nasal congestion, dermatological disorders such as psoriasis, allergies, Parkinson's disease, Alzheimer's disease, acute heart failure, angina disorders, delirium and dystonias such as Huntington's disease. They can also be used to regulate steroid hormone disorders, epinephrine release disorders, electrolyte balance disorders, endocrine disorders, memory disorders, somatosensory disorders, metabolic disorders, behavioural disorders, drug addiction, migraine, Addison's disease, Cushing's disease, prevent miscarriage, induce labour or to treat dysmenorrhoea. The present cDNA sequence encodes rat G-protein coupled receptor, SNORF72 related to the invention.

50 Sequence 1231 BP; 253 A; 367 C; 292 G; 319 T; 0 other;

Query Match	100.0%;	Score 1231;	DB 22;	Length 1231;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1231;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Oy	1	gctctgagatctttaaagctcagcagaaggggaaaaacttgaaaatgctctccgagctccagatcc	60
Db	1	gtctgtagatctttaaagccacagataaatgggaaaaacttgaaaatgctctcccgagctccagatcc	60
Oy	61	tctcaatgaagtaacttgaaacagacagagagagtaacttggcccaactgtgtgagaccagaagc	120
Db	61	tctctatgaaagtaacttgaaacagacagagagagtaacttggcccaactgtgtgagaccagaagc	120
Oy	121	cagtgacccaaccctcccgaggtgtctctgagccaaagctgacatctctctgtgtggggtaat	180
Db	121	cagtgacccaaccctcccgaggtgtctctgagccaaagctgacatctctctgtgtggggtaat	180
Oy	181	gggcaatctctctgtgtgagcagatgtgtgtccgcacatcagaacttgaagaacacccacaa	240
Db	181	gggcaatctctctgtgtgagcagatgtgtgtccgcacatcagaacttgaagaacacccacaa	240
Oy	241	ctactaatctctcagctctgagcagctcagaatctctgtctctcgtctctcttgaggatgacctt	300

Db	241	ctacatctctctcaagcttggcagatcctaaagctcgcgcgtccgcgtctcttggggaatgcctc	300
Qy	301	ggaaatctcagagatgttgcacaataacccctctccgtcttcggccgttvggagatgtactt	360
Db	301	ggaaatctcagagatgttgcacaataacccctctccgtcttcggccgttvggagatgtactt	360
Qy	361	caagagacgcccctcttcgaagacttgttgcttgcctccattcttcagtgctacacaggttag	420
Db	361	caagagacgcccctcttcgaagacttgttgcttgcctccattcttcagtgctacacaggttag	420
Qy	421	cgtaagagcctatgttggccatttggccaccccttcgcagccaagcttggaaagacgcggcg	480
Db	421	cgtaagagcctatgttggccatttggccaccccttcgcagccaagcttggaaagacgcggcg	480
Qy	481	acgggcctctcagatctctcagccttagctcggagctctctctgtgtgtcttcttcttgcaca	540
Db	481	acgggcctctcagatctctcagccttagctcggagctctctctgtgtgtcttcttcttgcaca	540
Qy	541	taccagatctcatgtgacataagttccagacattctccaaacgggtccctccgtaccctgc	600
Db	541	taccagatctcatgtgacataagttccagacattctccaaacgggtccctccgtaccctgc	600
Qy	601	agccacgtgcacagtcacacaaacccatgttgggtgtataacttgatcatccaaagtacacag	660
Db	601	agccacgtgcacagtcacacaaacccatgttgggtgtataacttgatcatccaaagtacacag	660
Qy	661	cttcctctctacaactccctccaatgacccctacacagctccctctactactcaatvgggct	720
Db	661	cttcctctctacaactccctccaatgacccctacacagctccctctactactcaatvgggct	720
Qy	721	caggctctgaagagaaatgaatccctctggaggagaaagctggcgtgtgaatatccaaagcc	780
Db	721	caggctctgaagagaaatgaatccctctggaggagaaagctggcgtgtgaatatccaaagcc	780
Qy	781	ctctagaagaatcagtcacacaagatgctgttctgttctgtccctcgltgttgcacatctgtg	840
Db	781	ctctagaagaatcagtcacacaagatgctgttctgttctgtccctcgltgttgcacatctgtg	840
Qy	841	gaccccccctccatgttgaacgggcctctctctccagctcttggagaaagtggacagaatccct	900
Db	841	gaccccccctccatgttgaacgggcctctctctccagctcttggagaaagtggacagaatccct	900
Qy	901	ggctgcgtgtgtcaacctcaatcctatgtgagatcaggtgtctctcttatactgaagctccg	960
Db	901	ggctgcgtgtgtcaacctcaatcctatgtgagatcaggtgtctctcttatactgaagctccg	960
Qy	961	gttcaaaccccatctataacacctcctgtcttcggcgcttcggcgcgcccttcgaatgt	1020
Db	961	gttcaaaccccatctataacacctcctgtcttcggcgcttcggcgcgcccttcgaatgt	1020
Qy	1021	tgctctccctctacccgcgaatgtgtgccaatcccggcgaatggcgaacagggacctccagccca	1080
Db	1021	tgctctccctctacccgcgaatgtgtgccaatcccggcgaatggcgaacagggacctccagccca	1080
Qy	1081	gaagatcactctcttgcacagaatgtcaacctcgtgtgggctctgacagagatgtcagcgcccca	1140
Db	1081	gaagatcactctcttgcacagaatgtcaacctcgtgtgggctctgacagagatgtcagcgcccca	1140
Qy	1141	gtccctgtgtcagtcattccatccacaacaccccttaccacgcccccttgcagagga	1200
Db	1141	gtccctgtgtcagtcattccatccacaacaccccttaccacgcccccttgcagagga	1200
Qy	1201	gttaccataaaagagtgtctcagaagcctc	1231
Db	1201	gttaccataaaagagtgtctcagaagcctc	1231

RESULT	2
AAH43078	
ID	AAH43078 standard; DNA; 1185 BP
XX	
AC	AAH43078;
XX	

DT 15-OCR-2001 (first entry)
XX Nucleotide sequence of a rat TGR-1 protein.
DE
XX TGR-1; neuromedin U; hypertension; stress disease; ss.
KM
XX Rattus sp.
OS
XX
FH Key location/qualifiers
FT 1..1185
FT /*tag= a
FT /product= "TGR-1"
FT /note= "no termination codon given"
XX
XX WO200157524-A1.
XX
XX 09-AUG-2001.
XX
XX 02-FEB-2001; 2001WO-JP00746.
XX
XX 04-FEB-2000; 2000JP-0032773.
PR 24-FEB-2000; 2000JP-0052252.
PR 30-MAR-2000; 2000JP-0097896.
PR 19-JUN-2000; 2000JP-0187536.
XX
XX (TAKE) TAKEDA CHEM IND LTD.
XX
XX Hinuma S, Shintani Y, Hosoya M, Fujii R, Moriya T, Matsui H;
P1 Okubo S;
XX
DR WPI: 2001-488917/53.
DR P-PSDB: AAG63367.
XX
XX Identifying predicted or actual structures of two or more members of a
PT chemical or physical library by mass spectrometry comprising
PT correlating molecular mass measurements of two or more members with a
PT shared chemical history -
XX
XX Claim 9; Page 91-92; 95pp; Japanese.
XX
XX The present sequence encodes a rat TGR-1 protein. The specification
CC describes a method of screening a compound, which is capable of binding
CC properties of neuromedin U to TGR-1. The method is useful for screening
CC preventative and remedies for hypertension, stress diseases, etc..
CC TGR-1 antagonists are also useful for treating the same diseases.
XX
XX Sequence 1185 BP; 242 A; 360 C; 277 G; 306 T; 0 other:
SO

Query Match 95.6%; Score 1177; DB 22; Length 1185;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1180; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 23 atgggaacttgaaatgctctcctgataccagatctctcctcgaatgaagtaacttgacagc 82
DB 1 atgggaacttgaaatgctctcctgataccagatctctcctcgaatgaagtaacttgacagc 60
OY 83 acagaggagtaacttgcccaactgtgtgacccaagcgagtgactatacccttcggg 142
DB 61 acagaggagtaacttgcccaactgtgtgacccaagcgagtgactatacccttcggg 120
OY 143 tctgtgcatatgagctatctcctctgtgtgaggaatgagcaactctgtgtgcatg 202
DB 121 tctgtgcatatgagctatctcctctgtgtgaggaatgagcaactctgtgtgcatg 180
OY 203 gtgattgtccagatcagacttgaaagacacccaactactatctcttcagcttgaca 262
DB 181 gtgattgtccagatcagacttgaaagacacccaactactatctcttcagcttgaca 240
OY 263 gtctcagatctgtgtctgtctgtgtgaggaatgagcaactcttcagagtgagcac 322
DB 241 gtctcagatctgtgtctgtctgtgtgaggaatgagcaactcttcagagtgagcac 300

OY 323 aataccttctcgttctcgagcgtgtgagatgactatccaagacagccctctcgagact 382
DB 301 aataccttctcgttctcgagcgtgtgagatgactatccaagacagccctctcgagact 360
OY 383 gtgtgtcttgcctcattctcagtgatgacacaggttgagcgtatgagcgtatgagcatt 442
DB 361 gtgtgtcttgcctcattctcagtgatgacacaggttgagcgtatgagcgtatgagcatt 420
OY 443 gtccaccccttccagagccaagctggaagacagcgagcgagcgcccttcagatctccagc 502
DB 421 gtccaccccttccagagccaagctggaagacagcgagcgagcgcccttcagatctccagc 480
OY 503 ctgactgtgagctctcgtgtgtctcttcttcttgcccaatacagatcctcagcctcag 562
DB 481 ctgactgtgagctctcgtgtgtctcttcttcttgcccaatacagatcctcagcctcag 540
OY 563 ttccagacattcccaacgagctcctcgttaacctgtgtcagccaacttgacagtcaccaa 622
DB 541 ttccagacattcccaacgagctcctcgttaacctgtgtcagccaacttgacagtcaccaa 600
OY 623 cccatgtgtgtgtataacttgatcattcaagctaccaaacttctcttcttatactctcca 682
DB 601 cccatgtgtgtgtataacttgatcattcaagctaccaaacttctcttcttatactctcca 660
OY 683 atgacccctcagcgtcctctactactcctcatgaggcgctcagagctgaagagatgaatcc 742
DB 661 atgacccctcagcgtcctctactactcctcatgaggcgctcagagctgaagagatgaatcc 720
OY 743 ctgtgagcgaacaaagtgtgtgtgaatattcacaagaccccttgaagaatgacagcaag 802
DB 721 ctgtgagcgaacaaagtgtgtgtgaatattcacaagaccccttgaagaatgacagcaag 780
OY 803 atgctgt 862
DB 781 atgctgt 840
OY 863 ctctctcagcttctgtgaagagtgagacagagtcctcgtgtgtgtgtgtgtgtgtgtgt 922
DB 841 ctctctcagcttctgtgaagagtgagacagagtcctcgtgtgtgtgtgtgtgtgtgtgt 900
OY 923 catgtgtatcagatgt 982
DB 901 catgtgtatcagatgt 960
OY 983 ctccgtctcgt 1042
DB 961 ctccgtctcgt 1020
OY 1043 tgcacatcccgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1102
DB 1021 tgcacatcccgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1080
OY 1103 tgtcacctcgt 1162
DB 1081 tgtcacctcgt 1140
OY 1163 cacaac 1207
DB 1141 cacaac 1185

RESULT 3
AAD08008
ID AAD08008 standard; cDNA; 1298 BP.
XX
XX AAD08008;
XX
XX 07-AUG-2001 (first entry)
XX
XX Human G-protein coupled receptor, SNORF72 cDNA.
DE
XX Human; G-protein coupled receptor; SNORF72; neuromedin U neuropeptide;
KW NMU; inflammation; arthritis; autoimmune disease; septicaemia; psychotic;

mental retardation; transplant rejection; neurological disorder; anxiety;
respiratory disorder; depression; schizophrenia; dementia; obesity; pain;
gastrointestinal disorder; hypertension; hypotension; epilepsy; diabetes;
ischemia; stroke; cancer; sexual disorder; circadian disorder; anorexia;
dermatological; psoriasis; Parkinson's disease; nausea; bulimia; allergy;
Alzheimer's disease; AIDS; hormonal disorder; memory disorder; migraine;
cardiovascular disorder; renal disorder; bone disease; delirium; asthma;
Cushing's disease; dysmenorrhoea; antidiagonal; cytostatic; osteoporosis;
metabolic disorder; behavioural disorder; Addison's disease; dyskinesia;
tranquilliser; antitumor; antidiabetic; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 27.1274

FT /tag= a

FT /product= "Human SNORF72 receptor"

XX MO200144297-A1.

XX 21-JUN-2001.

XX 13-DEC-2000; 2000MO-US33787.

XX 17-DEC-1999; 99US-0466435.

PR 25-APR-2000; 2000US-0558099.

PR 30-JUN-2000; 2000US-0609146.

XX (SYNA-) SYNAPTC PHARM CORP.

XX Bonini JA, Lerman GS, Quan Y, Ogozalek K;

XX MPI: 2001-390240/41.

DR P-PSDB: AAE03629.

PT A purified mammalian SNORF62 or SNORF72 receptor protein for

PT identification of compounds to treat e.g. inflammation, arthritis,

PT autoimmune diseases, transplant rejection, AIDS, cancer -

XX Claim 42: Fig 3: 256bp: English.

XX The invention relates to human G-protein coupled receptors, SNORF62 and

CC SNORF72 and their corresponding cDNA molecules. SNORF62 and SNORF72

CC receptors are specific for neuromedin U (NMU) neuropeptides, hence they

CC are also known as NMU receptors. The agonist and antagonist of NMU

CC receptors are useful for treating an abnormality in a subject that is

CC alleviated by decreasing or increasing the activity of NMU receptor.

CC The NMU receptors serve as a valuable tool for designing drugs which are

CC useful for treating various pathophysiological conditions such as

CC inflammation, arthritis, autoimmune diseases, transplant rejection,

CC graft vs host disease, bacterial, fungal, protozoan and viral infections,

CC septicemia, AIDS, pain, psychotic and neurological disorders, including

CC anxiety, depression, schizophrenia, dementia, mental retardation, memory

CC loss, epilepsy, neuromotor disorders, respiratory disorders, asthma,

CC eating/body weight disorders including obesity, bulimia, diabetes,

CC anorexia, nausea, hypertension, hypotension, vascular and cardiovascular

CC disorders, ischemia, stroke, cancers, sexual disorders, circadian

CC disorders, renal disorders, bone diseases including osteoporosis, benign

CC prostatic hypertrophy, gastrointestinal disorders, nasal congestion,

CC dermatological disorders such as psoriasis, allergies, Parkinson's

CC disease, Alzheimer's disease, acute heart failure, angina disorders,

CC delirium and dyskinesias such as Huntington's disease. They can also be

CC used to regulate steroid hormone disorders, epinephrine release

CC disorders, electrolyte balance disorders, endocrine disorders, memory

CC disorders, somatosensory disorders, metabolic disorders, behavioural

CC disorders, drug addiction, migraine, Addison's disease, Cushing's

CC disease, prevent miscarriage, induce labour or to treat dysmenorrhoea.

CC The present cDNA sequence encodes human G-protein coupled receptor,

CC SNORF72.

XX Sequence 1298 BP: 278 A; 405 C; 282 G; 333 T; 0 other;

50

Query Match 65.0%: Score 799.6; DB 22; Length 1298;
Best Local Similarity 79.2%: Pred. No. 3.2e-227;
Matches 968; Conservative 0; Mismatches 239; Indels 15; Gaps 1;

QY 5 tggatttaagctcagtaatggaagaacttgaataatgcttcctgatacc----- 55

DB 18 tggatttaatgctcaggaatggaagaacttcagaatgcttcctgataccagaaga 77

QY 56 -----gattctctatgaatgacttgaagcagcagagagacttggcccaatggt 109

DB 78 ctgaagaatcattccagaacaacctgacagcagcagagagatctggtcctcctcgc 137

QY 110 ggaaccaagcgaatgacattcccttcgcgtgtctgctgacatacgctatctctc 169

DB 138 ggaactcggcgcagcacttctctcccgctgtctgtggtgatactgccaatttttg 197

QY 170 gtgggggaatgagcaatctctctgtgcatgtgatactgctcgaacatcagaattgag 229

DB 198 gtgggggcatgtgcaatgctctgtgtgctgtgctgtgattctgcagcagcagctagag 257

QY 230 acaccaccaatactatctcttcaagcttgacagctcagaatctgctgtctctctg 289

DB 258 acgcccaccaatactactctcttccagctgtgctgtcctgacccctcctggtctcct 317

QY 290 ggaatgctctggaatactcagagatgtgacaaatcccttctctgctcgagcgtg 349

DB 318 ggaatgcccctggagatgtctatgagatgtgctgcgaactaccccttctctgtcgagccgtg 377

QY 350 ggaatgtaacttaagaacagccctctcagagatctgtgcttctgctccatctcagatgc 409

DB 378 ggtctacttaagaagcagcgcctctttagacggtgtctctgcctccatctcagatc 437

QY 410 accaaggttagctgagcagctatgtgacattgtgcaccccttctcagcagcaagctgag 469

DB 438 accacgcgtcaggtgagagcgtacgttgccacccacccgcttccgcccacaaatcag 497

QY 470 agcaagcgacagcgccctcagagatctcagcctcagctgagatctctctgtgctctt 529

DB 498 agcacccgagcgccgagccctcagagatctcagcctcagcgtctgtgctctcctc 557

QY 530 tcttggccaataacagacatccatggaatccagatctccacacttcccaaggtctctcc 589

DB 558 tcccgccacaacacagacatccatgcatcaagtctccactcttcccaatggtgctctg 617

QY 590 gtaactgtcagcagcacttgcagatcccaaacatggtggtgataacttgaatc 649

DB 618 gtcccaaggttgcgcacactgtaaggtacaaagcccaagtggaatctcaactcaatc 677

QY 650 caagctacagctctctctctctacatccctcccaatgacccatcagctcctctacac 709

DB 678 cagctacactctctctctctctctctctctccctcccaatgactgtcatagtctctacac 737

QY 710 ctcatggggtcagctgtaagaagagatgaatcccttgagcgcaacaagtgtgtgac 769

DB 738 ctcaagcagctcagctcgaagaagaacaatctcttgagcgagatggaaggaatgcaat 797

QY 770 attccaacacctcagaagaagtcacatcccaagatgctgtgtgtctgtctgtgt 829

DB 798 attcaaaagacctgagaanaacagtcacaaagaagatgctgtgtctgtctgtgt 857

QY 830 ggaactctgtgaccccttcattgtgacagcgctctctctcagcttctgtggaagatgg 889

DB 858 gctatctgttgagcccggttcccaatgacagcactctctcagcttctgtgagagaggg 917

QY 890 acagatccctgtgctgtgtgttcaacctcatalcctatgtgatacagtgctctcttatac 949

DB 918 agtgaatccctgtgctgtgttcaacctcagctcagctgtgtgtgtgtgtgtctctac 977

QY 950 ctgagctcgcggtcacaacctatctataaacctcctgtctgtgagctctccgagcgcc 1009

DB 978 ctgagctcagctgttcaaccttatactataacctactgtctgcgctctccagcgacga 1037

QY 1010 ttctgaaatgtgtctccctacctcgtcaaatgtgtgcatactcccgagcatcgccacagga 1069

Db 1038 ttccagaaatgagatctctctctccacaaacagtgccctcccaagacagaccacagttg 1097
Oy 1070 cctccagccaggaatcatctctctgacagaatgacactcgtgagctgacagagat 1129
Db 1098 ccactgcgcagcgaatacatctctctgacagaatgacactcgtgagctgacagagat 1157
Oy 1130 gcaggcccccagctccctcgtgacatcatccacaaacacacttaccagggccccc 1189
Db 1158 ataggctcccaatccctcgtgacatcatccacaaacacttaccacagccctc 1217
Oy 1190 tctgcagagagaggtacacaa 1211
Db 1218 tctagtgacacagatgtcaagaa 1239

RESULT 4
AAF80322
ID AAF80322 standard; DNA; 1658 BP.
XX
AC AAF80322;
XX
DT 29-JUN-2001 (first entry)
XX
DE Nucleotide sequence of a G-protein coupled receptor IGS4 long version.
XX
KW Human; G-protein coupled receptor; IGS4; IGS4A; IGS4B; schizophrenia;
KW nervous system disorder; psychiatric disorder; Parkinson's disease;
KW episodic paroxysmal anxiety; phobia; migraine; epilepsy; bulimia; stroke;
KW cardiovascular disease; heart failure; angina pectoris; obesity; emesis;
KW motility disorder; myocardial infarction; hypertension; dyslipidemia;
KW gastrointestinal disorder; inflammatory bowel disease; osteoporosis;
KW inflammation; infection; pain; cancer; immune disorder; allergy; sepsis;
KW gynecological disorder; ss.
XX
XX Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 55..1302
FT /tag="a
FT /product="G-protein coupled receptor IGS4A long version"
XX
XX MO200125269-A2.
XX
PD 12-APR-2001.
XX
XX 25-SEP-2000; 2000WO-EP09584.
XX
XX 24-SEP-1999; 99EP-0203140.
XX 24-SEP-1999; 99NL-1013140.
XX 28-JUL-2000; 2000EP-0202683.
XX 31-JUL-2000; 2000US-0222047.
XX
XX (SOLV) SOLVAY PHARM BV.
XX
XX Deleersnyder W, Berger C, Loeken C, Nys G, Venema J;
PI WPI: 2001-273568/28.
XX
XX P-PSDB; AAB67802.
XX
XX New G-protein coupled receptors and the polynucleotides encoding them,
XX useful for preventing, ameliorating or correcting nervous system
XX disorders, cardiovascular diseases, dyslipidemias, inflammations, pain
XX or cancers -
XX
XX Claim 1; Page 75-77; 102pp; English.
XX
XX The present sequence encodes the long version of a human G-protein
XX coupled receptor designated IGS4A. IGS4 exists in two polymorphic forms,
XX IGS4A and IGS4B. The IGS4 receptors and IGS4 polynucleotides are useful
XX for preventing, ameliorating or correcting dysfunction or diseases.
XX These diseases include peripheral nervous system, psychiatric and central
XX nervous system disorders (e.g. schizophrenia, episodic paroxysmal

CC anxiety, phobia, Parkinson's disease, migraine, epilepsy, bulimia or
CC stroke), cardiovascular diseases (e.g. heart failure, angina pectoris,
CC myocardial infarction or hypertension), dyslipidemias, obesity, emesis,
CC gastrointestinal disorders (e.g. inflammatory bowel disease or motility
CC disorders), osteoporosis, inflammations, infections (e.g. bacterial,
CC fungal, protozoan or viral), pain, cancers, immune disorders, allergies,
CC sepsis or gynecological disorders. Agonists or antagonists of IGS4 are
CC effective with regard to disorders of the nervous system, including the
CC central and peripheral nervous systems, disorders of the gastrointestinal
CC system, cardiovascular system, skeletal muscle, thyroid, lung or
CC genitourinary system, or immunological disease. The IGS4 polynucleotides
CC are useful as diagnostic reagents for detecting under-expression,
CC overexpression or altered expression of IGS4.
XX
XX Sequence 1658 BP; 377 A; 490 C; 346 G; 445 T; 0 other:
SQ

Query Match 65.0%; Score 799.6; DB 22; Length 1658;
Best Local Similarity 79.2%; Pred. No. 3.6e-227;
Matches 968; Conservative 0; Mismatches 239; Indels 15; Gaps 1;

Oy 5 tggatttaagctcagtaatgggaaacttgaataatgtcttcgtgacac----- 55
Db 46 tggatttaatgtcagagatggaataaacttcagaatgtcttcgtgacacagagaa 105
Oy 56 -----gatcctctatgaagtacttgaacagcacagagaggtacttggccactgtgt 109
Db 106 ctagaagaatccatcccaaaacacttgaacagcacagagaggtacttggccctctgc 165
Oy 110 ggaaccaagcagagtgacctatcccttcgcgtgtctgtgacgtatgcgtgtctctct 169
Db 166 ggaactcgcgcagacactctctctcccgctgtctgtgtgtgtatgtgtacatttttg 225
Oy 170 gtcggggttaagtggaactctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 229
Db 226 gtcgggggtatgtgcaatgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 285
Oy 230 aacaccacaactactatctctcagctgtgagctgtgacgtgtgtgtgtgtgtgtgtgtgt 289
Db 286 agccccaacactactactctctcagctgtgagctgtgtgtgtgtgtgtgtgtgtgtgtgt 345
Oy 290 ggaatgcctcttgaactcagacagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 349
Db 346 ggaatgcctcttgaactcagacagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 405
Oy 350 ggaatgcctcttgaactcagacagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 409
Db 406 ggtgtactactaagaagcgcctctcttgaagaccgtgtgtgtgtgtgtgtgtgtgtgtgt 465
Oy 410 accaaggttaggttagagctatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 469
Db 466 accaaggttaggttagagctatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 525
Oy 470 agcagcgcgcagcgcgcctcagacatccctcagctgtgtgtgtgtgtgtgtgtgtgtgtgt 529
Db 526 agcagcgcgcgcgcgcctcagacatccctcagctgtgtgtgtgtgtgtgtgtgtgtgtgt 585
Oy 530 tcttggccaataccagatcctcagacatccctcagctgtgtgtgtgtgtgtgtgtgtgtgt 589
Db 586 tccctgccaacacacagatcctcagacatccctcagctgtgtgtgtgtgtgtgtgtgtgtgt 645
Oy 590 gtaactgtgtcagcagcgcgcagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 649
Db 646 gtccaggttgcgcacccgt 705
Oy 650 caagctacagcttctctctctacatccctcagacatccctcagctgtgtgtgtgtgtgtgt 709
Db 706 cagtgatcct 765
Oy 710 ctcatgggtcctcagctgaagagatgaatcccttgaagcgaacaaagtgtgtgtgtgtgt 769
Db 766 ctcatgctcactcagctaaagaagaacatcttgaagcagatgaaggaagtgaat 825


```

Db 526 agcaccgagcgccgagccctcagaagtcctcgcagtcgtctggggtcttcctcgtcttc 585
Oy 530 tcttggccaataacagcatccatgagcaatcgaagtcagcagcattcccaacgggtctcc 589
Db 536 tcctggccaataacagcatccatgagcaatcgaagtcagcagcattcccaacgggtctcc 645
Oy 550 gtaccctgctcagcagcagtcagcaagtcacaaacccatgtgggtgtataactgtatca 649
Db 646 gtcccaaggtctcgcagcagcagtcagcaagtcacaaacccatgtgggtgtataactgtatca 705
Oy 650 caagctacagcagcagtcctcttctacatccctcccaatgagccctcaatcagcgtctctac 709
Db 706 caggtacagcagcagtcctcttctacatccctcccaatgagcagtcagtcagtcctctac 765
Oy 720 ctcaatggggtcagcagcagtcagcaagtcagcaagtcagcagtcagtcagtcagtcagtc 769
Db 766 ctcaatggggtcagcagcagtcagcaagtcagcaagtcagcagtcagtcagtcagtcagtc 825
Oy 770 attcagaagcagcagtcagcaagtcagcaagtcagtcagtcagtcagtcagtcagtcagtc 829
Db 826 attcagaagcagcagtcagcaagtcagcaagtcagtcagtcagtcagtcagtcagtcagtc 885
Oy 830 gcaatctgctgagcagcagtcagcaagtcagcaagtcagtcagtcagtcagtcagtcagtc 889
Db 886 gcaatctgctgagcagcagtcagcaagtcagcaagtcagtcagtcagtcagtcagtcagtc 945
Oy 890 acagagtcagcagcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 949
Db 946 agtgaatccctgagcagcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 1005
Oy 950 ctgagctcagcagcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 1009
Db 1006 ctgagctcagcagcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 1065
Oy 1010 ttgcaaatgtgtctcctcagcagcagtcagtcagtcagtcagtcagtcagtcagtcagtc 1069
Db 1066 ttgcaaatgtgtctcctcagcagcagtcagtcagtcagtcagtcagtcagtcagtcagtc 1125
Oy 1070 cctcagaagcagcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 1129
Db 1126 ccaactgagcagcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 1185
Oy 1130 gcaagcagcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 1189
Db 1186 ataggtcccaatccatgtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 1245
Oy 1190 tctgcaagagaggtacataaa 1211
Db 1246 tctagtgacaagatgtcaagaa 1267

```

RESULT 6

AAF80324 AAF80324 standard; DNA: 1658 BP.

AAF80324:

29-JUN-2001 (first entry)

Nucleotide sequence of a G-protein coupled receptor IGS4B long version.

Human; G-protein coupled receptor; IGS4; IGS4A; IGS4B; schizophrenia;
 nervous system disorder; psychiatric disorder; Parkinson's disease;
 episodic paroxysmal anxiety; phobia; migraine; epilepsy; bulimia; stroke;
 cardiovascular disease; heart failure; angina pectoris; obesity; emesis;
 motility disorder; myocardial infarction; hypertension; dyslipidemia;
 gastrointestinal disorder; inflammatory bowel disease; osteoporosis;
 inflammation; infection; pain; cancer; immune disorder; allergy; sepsis;
 gynecological disorder; ss.

Hom sapiens.
 Location/Qualifiers

Key

```

FT CDS 55..1302
FT /tag=A
FT /product="G-protein coupled receptor IGS4B long version"
FM
XX
FM W0200125269-A2.
XX
PD 12-APR-2001.
XX
PF 25-SEP-2000; 2000WO-EP09584.
XX
PR 24-SEP-1999; 99EP-0203140.
PR 24-SEP-1999; 99NL-1013140.
PR 28-JUL-2000; 2000EP-0202683.
PR 31-JUL-2000; 2000US-0222047.
XX
PA (SOLV ) SOLVAY PHARM BV.
XX
PI Deleersnyder W, Berger C, Loeken C, Nys G, Venema J.
XX
DR WPI: 2001-273568/28.
XX
DR P-PSDB; AAB67804.
XX
PT New G-protein coupled receptors and the polynucleotides encoding them,
PT useful for preventing, ameliorating or correcting nervous system
PT disorders, cardiovascular diseases, dyslipidemias, inflammations, pain
PT or cancers
XX
PS Claim 1: Page 83-85; 102pp; English.
XX
CC The present sequence encodes the long version of a human G-protein
CC coupled receptor designated IGS4B. IGS4 exists in two polymorphic forms,
CC IGS4A and IGS4B. The IGS4 receptors and IGS4 polynucleotides are useful
CC for preventing, ameliorating or correcting dysfunctions or diseases.
CC These diseases include peripheral nervous system, psychiatric and central
CC nervous system disorders (e.g. schizophrenia, episodic paroxysmal
CC anxiety, phobia, Parkinson's disease, migraine, epilepsy, bulimia or
CC stroke), cardiovascular diseases (e.g. heart failure, angina pectoris,
CC myocardial infarction or hypertension), dyslipidemias, obesity, emesis,
CC gastrointestinal disorders (e.g. inflammatory bowel disease or motility
CC disorders), osteoporosis, inflammations, infections (e.g. bacterial,
CC fungal, protozoan or viral), pain, cancers, immune disorders, allergies,
CC sepsis or gynecological disorders. Agonists or antagonists of IGS4 are
CC effective with regard to disorders of the nervous system, including the
CC central and peripheral nervous systems, disorders of the gastrointestinal
CC system, cardiovascular system, skeletal muscle, thyroid, lung or
CC genitourinary system, or immunological disease. The IGS4 polynucleotides
CC are useful as diagnostic reagents for detecting under-expression,
CC overexpression or altered expression of IGS4.
XX
SO Sequence 1658 BP; 378 A; 489 C; 346 G; 445 T; 0 other.

```

Query Match 64.7%; Score 796.4; DB 22; Length 1658;
 Best Local Similarity 79.1%; Pred. No. 3.2e-226;
 Matches 966; Conservative 0; Mismatches 241; Indels 15; Gaps 1;

```

Oy 5 tggatttaagctcaatgaggaacttgaaatgctctcgtgatac----- 55
Db 46 tggatttaatgctcagggatgaaacttcagaaatgctctcgtgatacagaagaa 105
Oy 56 -----gatcctcatgaagtactgaacacagacagaggaatgagccactgtat 109
Db 106 ctagaagatcctctcagaaacacactgaaacagacagaggaatgagccactgtat 165
Oy 110 gaccaaagcagcagtcagcagtcagtcagtcagtcagtcagtcagtcagtcagtc 169
Db 166 ggaactcgcgagcagcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 225
Oy 170 gtgggggtcaatgagcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 229
Db 226 gtgggggtcaatgagcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 285
Oy 230 acaccacaactatactctcagcttgagcagtcagcagtcagtcagtcagtcagtc 289

```

Dh	286	agcgcccaacaactctactctctcccttcagcgtcgtgcgtctctcgactccctgctcgtctctct	345
Oy	280	ggagtcgctctvgaaatctctacagagatgctgacacattaccccttcctcgttcggcgctg	349
Dh	346	ggaaatgcctccctgvgggtgctctatvgaaatgctggcgcaaatcaacccttcctctgctggcgcg	405
Oy	350	ggatgctactctcaagaagacgccccttcggaaagctgctgtgttcgttgcctccatctaaagtgc	409
Dh	406	ggctgctactctcaagaagagcgccctcttgaaacgcgtgtgcttcgctccatcctcaagctc	465
Oy	410	accacggtgactcgtatagagcgctatgtgcccattgtccacccttcctcgaaagactgag	469
Dh	466	accacccgcgtacgtgtaggcgtactgctgcccattccacccgtctccgcgcgaactgcag	525
Oy	470	agcacgcggcgagcgagcgccctccagaatcccttcagctcgtatgttgagactctctcgtgtctt	529
Dh	526	agcacccgcgcgcggcgccctccagaatcccttcgcatccgtctcgtgggctctcccgctcttc	585
Oy	530	tccttcggcccaatlaacagactcctcaatgcatccatcttcacagaccttcttcacacggtccctcc	589
Dh	586	tcctctggcccaacacacagcatccatcgaatgcatccactctccccaatggcgctcgt	645
Oy	590	gtacctggtctcagccaccctgcagatccacaacccaatgctgggtgtataactgtatc	649
Dh	646	gtcccaaggtcttgcgccaccctgtacggtctatcaaaagcccaatgtgactctacattctacatc	705
Oy	650	caagctacacagctctccctctctctacatccctccaaatgaacctccatcaagcgctccctacac	709
Dh	706	caggtcacctcctctccatctacatccctccccaatgacgtcgtacagtgtccctccatcac	765
Oy	710	ctcacatgggctcagagctgaaagagatgaaatccctctgaagcgcaacaaagtgtgctgact	769
Dh	766	ctcatgycatcagagctaaagaagaacaaatctctcttgagcgagatgaaaggaatgcaaat	825
Oy	770	attcaacagaccctctaaagaagctcagatccacacagatgctgtgtctctgctcgtctgtct	829
Dh	826	attcaaaagaccctgcgaataaaatccagttcaaaagaatgctgtctgtctctgacttaagtct	885
Oy	830	ggcaatctgctgaaacccctctccatgtgaaacgcgctctctctcaagctcttgaaagaagctg	889
Dh	886	gctactctgttgggcccgctcccaatgacagcgtctctctcaagctcttgtagagagagtg	945
Oy	890	acaaagatccctgctgctgtgtgtccaaacctcatccatgtgtatcaaggtctctcttcat	949
Dh	946	actgaatccctggtgtgtgtgtctcaacctgtccatctgctcaatgtgtctcaagtctctactctac	1005
Oy	950	ctgagctcccggtgcacaaaccccatatctataaaccctctctccggcggtctccggcgcgcc	1009
Dh	1006	ctggagctcagctgtgcacaaaccccatatctataaaccctgtctccggcgtctccaggcagca	1065
Oy	1010	tttcgaaatgtgtctccctcctaacctgcgaatgggtgcatacccgcgatctgcgcacagagga	1069
Dh	1066	tctcaagaatgtgactctctctctctccacaacacagtgcatctccagaatgaaccacaatg	1125
Oy	1070	gctccaaagcccaagaagatcatctctcttgacaaagatgctcacctggtgagatgtgacagaagt	1129
Dh	1126	ccaactgcgcacggaaacatctctccgtgacagaatgccaactgtgtgagatgcgcgcaagat	1185
Oy	1130	gcagagcccccaggttccctctgtctgactcatccatccacaacaaacaaactcttcacacgcccc	1189
Dh	1186	ataggtccccaattctctatgtctcaagtatccgtgcgacaaatctccaactcccaacagccctc	1245
Oy	1190	tgtgcagagagaggtacacataa	1211
Dh	1246	tctatgtgaaacagatgtccaagaa	1267

XX		
DT	29-JUN-2001	(first entry)
DE	Nucleotide sequence of G-protein coupled receptor IGS4B short version.	
DD		
XX		
KW	Human; G-protein coupled receptor; IGS4; IGS4A; IGS4B; schizophrenia;	
KM	nervous system disorder; psychiatric disorder; Parkinson's disease;	
KM	episodic paroxysmal anxiety; phobia; migraine; epilepsy; bulimia; stroke;	
KM	cardiovascular disease; heart failure; angina pectoris; obesity; emesis;	
KM	mottility disorder; myocardial infarction; hypertension; dyslipidemia;	
KM	gastrointestinal disorder; inflammatory bowel disease; osteoporosis;	
KM	inflammation; infection; pain; cancer; immune disorder; allergy; sepsis;	
KM	gynecological disorder; ss.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	64..1302
FT		/tag= "A
FT		/product= "G-protein coupled receptor IGS4B short
FT		version"
PN	WO200125269-A2.	
XX		
PD	12-APR-2001.	
XX		
PF	25-SEP-2000; 2000WO-EP09584.	
PR	24-SEP-1999; 99EP-0203140.	
PR	24-SEP-1999; 99NL-1013140.	
PR	28-JUL-2000; 2000EP-0202683.	
PR	31-JUL-2000; 2000US-0222047.	
XX		
PA	(SOLV) SOLVAY PHARM BV.	
PI	Deleersnijder W, Berger C, Loeken C, Nys G, Venema J;	
PI	WPI: 2001-273568/28.	
DR	P-Patent: AAB67805.	
XX		
PT	New G-protein coupled receptors and the polynucleotides encoding them,	
PT	useful for preventing, ameliorating or correcting nervous system	
PT	disorders, cardiovascular diseases, dyslipidemias, inflammations, pain	
PT	or cancers .	
PS	Claim 1; page 87-89; 102pp; English.	
XX		
CC	The present sequence encodes the short version of a human G-protein	
CC	coupled receptor designated IGS4B. IGS4 exists in two polymorphic forms,	
CC	IGS4A and IGS4B. The IGS4 receptors and IGS4 polynucleotides are useful	
CC	for preventing, ameliorating or correcting dysfunctions or diseases.	
CC	These diseases include peripheral nervous system, psychiatric and central	
CC	nervous system disorders (e.g. schizophrenia, episodic paroxysmal	
CC	anxiety, phobia, Parkinson's disease, migraine, epilepsy, bulimia or	
CC	stroke), cardiovascular diseases (e.g. heart failure, angina pectoris,	
CC	myocardial infarction or hypertension), dyslipidemias, obesity, emesis,	
CC	gastrointestinal disorders (e.g. inflammatory bowel disease or mottily	
CC	disorders), osteoporosis, inflammations, infections (e.g. bacterial,	
CC	fungal, protozoan or viral), pain, cancers, immune disorders, allergies,	
CC	sepsis or gynecological disorders. Agonists or antagonists of IGS4 are	
CC	effective with regard to disorders of the nervous system, including the	
CC	central and peripheral nervous systems, disorders of the gastrointestinal	
CC	system, cardiovascular system, skeletal muscle, thyroid, lung or	
CC	genitourinary system, or immunological disease. The IGS4 polynucleotides	
CC	are useful as diagnostic reagents for detecting under-expression,	
CC	overexpression or altered expression of IGS4.	
XX		
SQ	Sequence 1658 BP; 378 A; 489 C; 346 G; 445 T; 0 other:	

```

OY 5 tggatttaagctcgaatggaagaacttgaaaatgcttcctggtatccac----- 55
DB 46 tggatttaagctcgaatggaagaacttgaaaatgcttcctggtatccacagaagaa 105
OY 56 -----gatcctctcaatgaatgacttgaacagacagagagatcttggccacgtgt 109
DB 106 ctagaagaatccatccagaagaacacttgaacagacagagagatcttggccacgtgt 165
OY 110 ggaaccaagcgagtgacatccctccctggtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 169
DB 166 ggaacctcgcgagacacatctctccctccctggtgtgtgtgtgtgtgtgtgtgtgtgt 225
OY 170 gttggggaatggaatcctctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 229
DB 226 gttggggaatggaatcctctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 285
OY 230 acaaccaacaactatctctctcagcttgcagtcacagatctgtgtgtgtgtgtgtgt 289
DB 286 acgccaacaacaactatctctctcagcttgcagtcacagatctgtgtgtgtgtgtgtgt 345
OY 290 ggaatgctcttgaatctcagaatggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 349
DB 346 ggaatgctcttgaatctcagaatggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 405
OY 350 ggaatgctcttgaatctcagaatggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 409
DB 406 ggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 465
OY 410 accaagtgatcgaatgagcgtatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 469
DB 466 accaagtgatcgaatgagcgtatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 525
OY 470 agcaagcgcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgag 529
DB 526 agcaagcgcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgag 585
OY 530 tcttggccaataacagatccatccatccatccatccatccatccatccatccatccat 589
DB 586 tcttggccaataacagatccatccatccatccatccatccatccatccatccatccat 645
OY 590 gtaacctgtcagacacacacacacacacacacacacacacacacacacacacacac 649
DB 646 gtaacctgtcagacacacacacacacacacacacacacacacacacacacacacac 705
OY 650 caagctacacagcttctctctctctctctctctctctctctctctctctctctctct 709
DB 706 caagctacacagcttctctctctctctctctctctctctctctctctctctctctct 765
OY 710 ctcaatggcgctcagacgtgaagagagatgaaatccttgaagcgagaaagtggtgtgaat 769
DB 766 ctcaatggcgctcagacgtgaagagagatgaaatccttgaagcgagaaagtggtgtgaat 825
OY 770 attcaagaacccctagaagaatcagtcacacacacacacacacacacacacacacac 829
DB 826 attcaagaacccctagaagaatcagtcacacacacacacacacacacacacacacac 885
OY 830 ggaactgtgtgaacccctcctcctcctcctcctcctcctcctcctcctcctcctcct 889
DB 886 ggaactgtgtgaacccctcctcctcctcctcctcctcctcctcctcctcctcctcct 945
OY 890 acaagatccctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 949
DB 946 actgaatccctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1005
OY 950 ctgagctcgcggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1009
DB 1006 ctgagctcgcggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1065
OY 1010 ttctgaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1069
DB 1066 ttctgaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1125

```

```

OY 1070 cctcagccccaagaatcatctcttgaacgaatgtacacctcgttggagctgacagaagat 1129
DB 1126 ccactcgcagcggaacatctctctgacgaatgtacacctcgttggagctgacagaagat 1185
OY 1130 gcaagcccccagctcctcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1189
DB 1186 atagtcaccaatctctatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1245
OY 1190 tgtcaggaagaggtaccataa 1211
DB 1246 tctagtgacagatgtcaagaa 1267

RESULT 8
AAZ33297
ID AAZ33297 standard; cDNA: 1360 BP.
XX
XX AAZ33297;
AC
XX
XX 21-FEB-2000 (first entry)
DT
XX
XX Human neurotensin-like receptor encoding cDNA.
DE
XX
XX Human; neurotensin-like receptor; NLR: G-protein coupled receptor;
KW central nervous system; anesthesia; analgesia; ds.
XX
XX Homo sapiens.
OS
XX W09955732-A1.
PN
XX 04-NOV-1999.
PD
XX 15-APR-1999; 99WO-SE00598.
PE
XX 24-APR-1998; 98SE-0001455.
PR
XX (ASTR-) ASTRA PHARMA INC.
PA (ASTR) ASTRA AB.
PI
XX Ahmad S, Cao J, O'Donnell D, Walker P;
XX WPI: 2000-052803/04.
DR P-PSDB: AAY5292.
XX
XX Novel neurotensin-like receptor, useful for identifying agents for
PT producing anesthesia or analgesia
PI
XX
XX Claim 11: Fig 1; 46pp; English.
PS
XX
XX The present sequence encodes a human G-protein coupled receptor
CC neurotensin-like receptor (NLR). The NLR polynucleotide and protein
CC can be used to isolate compounds that bind, (ant)agonize or alter
CC the activity or expression of the NLR. The NLR is a G-protein coupled
CC receptor which is expressed in the central nervous system and shares
CC homology with human neurotensin receptor. The receptors can be used in
CC assays to identify agents for producing anesthesia and analgesia.
XX
XX
XX Sequence 1360 BP; 293 A; 423 C; 294 G; 350 T; 0 other:
SQ

Query Match 64.3%; Score 791.6; DB 21; Length 1360;
Best Local Similarity 78.8%; Pred. No. 7.8e-225;
Matches 963; Conservative 0; Mismatches 244; Indels 15; Gaps 1;
OY 5 tggatttaagctcgaatggaagaacttgaaaatgcttcctggtatccac----- 55
DB 56 tggatttaagctcgaatggaagaacttgaaaatgcttcctggtatccacagaagaa 115
DB 106 ctagaagaatccatccagaagaacacttgaacagacagagagatcttggccacgtgt 165
OY 56 -----gatcctctcaatgaatgacttgaacagacagagagatcttggccacgtgt 109
DB 116 ctagaagaatccatccagaagaacacttgaacagacagagagatcttggccacgtgt 175
OY 110 ggaaccaagcgagtgacatccctccctcctcctcctcctcctcctcctcctcctcct 169

```



```

Db 176 ggaacctcgagcagcaactcttcccccgtctgtgtatgtcgaattttgtc 235
170 gtgggggtataggcaattcttctgtgtcagatgtgtccgaataagacttgaag 229
236 gtgggggtataggcaattcttctgtgtcagatgtgtccgaataagacttgaag 295
230 aacacacacacacacacacacacacacacacacacacacacacacacacacac 289
296 acgcccacacacacacacacacacacacacacacacacacacacacacacacac 355
290 gggatgctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 349
356 ggaatgccccctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 415
350 ggaatgctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 409
416 ggtctgtctatcaagaagcgcctcttctgtgtgtgtgtgtgtgtgtgtgtgt 475
410 accacggtatagcgtatagcgtatgtgtgtgtgtgtgtgtgtgtgtgtgtgt 469
476 accacggtatagcgtatagcgtatgtgtgtgtgtgtgtgtgtgtgtgtgtgt 535
470 agcagcggcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 529
536 agcagcggcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 595
530 tcttggccaataacacacacacacacacacacacacacacacacacacacac 589
596 tccctgtcccaacacacacacacacacacacacacacacacacacacacac 655
590 gtacctgtgtcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 649
656 gtcccaaggtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 715
650 caagctacacagcttctcttctacatctcccaatgaacctcaacagctccctac 709
716 cagctacacctctctctacatctcccaatgaacctcaacagctccctac 775
710 cctacagcgtcagcgtcagcgtcagcgtcagcgtcagcgtcagcgtcagcgt 769
776 cctacagcgtcagcgtcagcgtcagcgtcagcgtcagcgtcagcgtcagcgt 835
770 attcacaagacccctcagaagaagtcagtcacaagaatgtgtgtgtgtgtgtgt 829
836 attcacaagacccctcagaagaatgtgtgtgtgtgtgtgtgtgtgtgtgtgt 895
830 gccatctgtcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 889
896 gctatctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 955
890 aagaagctcctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 949
956 actgaatccctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1015
950 ctgaagctcctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1009
1016 ctgaagctcctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1075
1010 ttctgaaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1069
1076 ttctgaaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1135
1070 cctcagaagcagaagaatctcttctcagaagaatctcttctcagaagaatct 1129
1136 ccactgtccagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1195
1130 gcaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1189
1196 aataggtcccaatctcctatagtcagtcagtcagtcagtcagtcagtcagtc 1255
1190 tgtcagaagaggtacacataa 1211

```

```

Db 1256 tctagtgaacagatgtcaagaa 1277

RESULT 9
AAH43072
ID AAH43072 standard; DNA; 1245 BP.
XX
AC AAH43072;
XX
DE 15-OCT-2001 (first entry)
XX
DE Nucleotide sequence of a human TGR-1 protein.
XX
DE TGR-1; neuromedin U; hypertension; stress disease; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1245
FT FT /*tag= a
FT FT /product= "TGR-1"
FT FT /note= "no termination codon given"
XX
PN WO200157524-A1.
XX
PD 09-AUG-2001.
XX
PF 02-FEB-2001; 2001WO-UP00746.
XX
PR 04-FEB-2000; 2000JP-0032773.
PR 24-FEB-2000; 2000JP-0052252.
PR 30-MAR-2000; 2000JP-0097896.
PR 19-JUN-2000; 2000JP-0187536.
XX
PA (TAKE ) TAKEDA CHEM IND LTD.
XX
PI Hinuma S, Shintani Y, Hosoya M, Fujii R, Moriya T, Matsui H;
PI Okubo S;
XX
DR WPI: 2001-488917/53.
DR P-PSDB: AM63353.
XX
PT Identifying predicted or actual structures of two or more members of a
PT chemical or physical library by mass spectrometry comparing
PT correlating molecular mass measurements of two or more members with a
PT shared chemical history -
XX
PS Claim 9; Page 79-80; 95pp; Japanese.
XX
CC The present sequence encodes a human TGR-1 protein. The specification
CC describes a method of screening a compound, which is capable of binding
CC properties of neuromedin U to TGR-1. The method is useful for screening
CC preventative and remedies for hypertension, stress diseases, etc..
CC TGR-1 antagonists are also useful for treating the same diseases.
XX
SQ Sequence 1245 BP; 268 A; 393 C; 268 G; 316 T; 0 other;

Query Match 64.2%; Score 790.8; DB 22; Length 1245;
Best Local Similarity 79.2%; Pred. No. 1.3e-224;
Matches 958; Conservative 0; Mismatches 237; Indels 15; Gaps 1;

Db 17 tcaagtaatgggaatctgaagaatgtctctctgattccac-----gattct 61
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
Db 4 tcaaggaatgggaatctgaagaatgtctctctgattccacagaagaatgaatcca 63
62 cctcagaatctgtgacagacagaggaatgtgtgtgtgtgtgtgtgtgtgtgtgtgt 121
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
Db 64 ttcagaagaacactgtgaacagcagcagaggtatctgtgtgtgtgtgtgtgtgtgt 123
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
Qy 122 agtgaactatccctcctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 181
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
Db 124 agcacttctctcctcccggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 183

```


Oy	182	ggcaatcttcgagtgatcagatgttcttcgcagcatcagactttgaagacacccccac	241
Db	184	ggcaatcttcgagtgatcagatgttcttcgcagcatcagactttgaagacacccccac	243
Oy	242	tactactctctcagctcgtgacgtccagatccgtctgtctcgtctctctctgtgagatgtccttg	301
Db	244	tactactctctcagctcgtgacgtctcgtctcgtctcgtctcgtctcgtctcgtctcgtctcgt	303
Oy	302	gaaatctcaagagatgtgccaattacaccttccctctcgttcgggcctctgtgatatctc	361
Db	304	gaaatctcaagagatgtgccaattacaccttccctctcgttcgggcctctgtgatatctc	363
Oy	362	aagacagcccccttcgagactctgtctgtctgtccatctcagatctcagacccgtttacg	421
Db	364	aagacagcccccttcgagactctgtctgtctgtccatctcagatctcagacccgtttacg	423
Oy	422	gttagagcgtatctgtgccaattgtccaacctctccagagccaagctgtgagagacagcgcgga	481
Db	424	gttagagcgtatctgtgccaattgtccaacctctccagagccaagctgtgagagacagcgcgga	483
Oy	482	cgggcctcccaagatccctcagcctcagctcttggaagtctctcgtgtctctctctcttcggacat	541
Db	484	cgggcctcccaagatccctcagcctcagctcttggaagtctctcgtgtctctctctcttcggacat	543
Oy	542	accgacatcccatctgagcatcaaggtctccagacattccccaagcggtcctccgttaacctgtcat	601
Db	544	accgacatcccatctgagcatcaaggtctccagacattccccaagcggtcctccgttaacctgtcat	603
Oy	602	gccacccctcagatccacccaacccaatgtgtgtgtatatacttgatctcatcagaatcagacg	661
Db	604	gccacccctcagatccacccaacccaatgtgtgtgtatatacttgatctcatcagaatcagacg	663
Oy	662	ttccctctctcatcatctccccaatgagacctcatagcgtctctcatcatcctcatctgagacgtc	721
Db	664	ttccctctctcatcatctccccaatgagacctcatagcgtctctcatcatcctcatctgagacgtc	723
Oy	722	aggtctgagagagatgaaatccctctgagcgagacaagatggtctgtgaaattccaagaccc	781
Db	724	aggtctgagagagagatgaaatccctctgagcgagacaagagaaatgaaatctccaagaccc	783
Oy	782	ctctgaagatcagatccacaagaatgctgtgtgtctgtctcgtctgtctgtctgtctgtctgt	841
Db	784	ctctgaagatcagatccacaagaatgctgtgtgtctgtctcgtctgtctgtctgtctgtctgt	843
Oy	842	accacctctcatctgtgacccggtctctcttcagctctctgtgagagatgtgacaagatcccttg	901
Db	844	gcccggtctccaatctgacccgactctcttcagctctctgtgagagatgtgacaagatcccttg	903
Oy	902	gctgctgtgttcaaacctcatcatctgataatgtataggtgtctctttatctgtggtcccg	961
Db	904	gctgctgtgttcaaacctctgcgtccaaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	963
Oy	962	gtcaaacccaattatcatctaaacctctgtctccagcggtctccggcggtcctctctgaaatgtct	1021
Db	964	gtcaaacccaattatcatctaaacctctgtctccagcggtctccggcggtcctctctgaaatgtct	1022
Oy	1022	gtctccacctatctgcagaaatgtgtgcatactcccggtcatccggccaacagagatccctcagccag	1081
Db	1024	atctctctcttcccaacaagaatgtgacatctcccaatgagccaacagaatctgtccacctctgcag	1083
Oy	1082	aagatctactctctgtgacagaatgttcaacctgtgtgaaatctgacagaagatgtcagcccccg	1141
Db	1084	cggaaatctctctgtgacagaatgtcgaatctgtgtgaaatctgacagaagatgtcagcccccg	1144
Oy	1142	ttccctctgtctcatctcatccacaacaacaacattcaacacgggtccctctgtcagagagag	1201
Db	1144	ttcccatctgtctcatctcatccacaacattccacctccacaacagcctctctgtctagtgacag	1203
Oy	1202	gtaccataaa 1211	
Db	1204	atgtcaagaa 1213	

RESULT	10	
ADAD01123	standard; cDNA: 1248 BP.	
ADAD01123		
ADAD01123		
02-NOV-2000	(first entry)	
Human orphan G protein-coupled receptor hrup6 cDNA.		
Human; orphan G protein-coupled receptor; GPCR; hrup6; drug screening; transmembrane receptor; signal cascade; ss.		
Homo sapiens.		
Key	Location/Qualifiers	
CDS	1..1248	
FT	/*tag= a	
FT	/product= "hrup6"	
FT	/note= "Human orphan G protein-coupled receptor"	
MO2000031258-A2.		
02-JUN-2000.		
13-OCT-1999;	99WO-US23687.	
20-NOV-1998;	98US-0109213.	
16-FEB-1999;	99US-0120416.	
26-FEB-1999;	99US-0121852.	
12-MAR-1999;	99US-0123946.	
12-MAR-1999;	99US-0123949.	
28-MAY-1999;	99US-0136436.	
28-MAY-1999;	99US-0136437.	
28-MAY-1999;	99US-0136439.	
28-MAY-1999;	99US-0136567.	
28-MAY-1999;	99US-0137127.	
28-MAY-1999;	99US-0137131.	
29-JUN-1999;	99US-0141448.	
29-SEP-1999;	99US-0156555.	
29-SEP-1999;	99US-0156633.	
29-SEP-1999;	99US-0156634.	
01-OCT-1999;	99US-0157280.	
01-OCT-1999;	99US-0157281.	
01-OCT-1999;	99US-0157282.	
01-OCT-1999;	99US-0157293.	
12-OCT-1999;	99US-0157294.	
12-OCT-1999;	99US-0416760.	
12-OCT-1999;	99US-0417044.	
(AREN-) ARENA PHARM INC.		
Chen R, Dang HT, Liaw CW, Lin I;		
WPI: 2000-400068/34.		
P-PSDB: AAY71296.		
Novel human orphan G protein-coupled receptors and the encoding cDNAs for use in the identification of G protein-coupled receptor agonists -		
Claim 21; Page 56-57; 102pp; English.		
The present sequence is a cDNA encoding hrup6, an endogenous human orphan G protein-coupled receptor (GPCR). The full length hrup6 cDNA was cloned by RT-PCR using human thymus cDNA as template.		
The orphan GPCR of the invention, like all GPCRs has seven transmembrane alpha helices with an extracellular N-terminus and an intracellular C-terminus. However, no endogenous ligands has yet been identified for the proteins of the invention. The orphan GPCRs may be used in the identification of their endogenous ligands, and to screen potential GPCR agonists and antagonists for use as pharmaceutical agents. The proteins		

CC may also be used in the study of GPCR-mediated signalling cascades, and
CC to elucidate their precise role in normal and diseased human conditions.
CC Nucleic acid encoding human orphan GPCRs may be used for tissue
CC localisation expression analysis to provide information about their
CC function in healthy and pathological states.

XX Sequence 1248 BP; 269 A; 393 C; 269 G; 317 T; 0 other;

Query Match 64.2%; Score 790.8; DB 21; Length 1248;
Best Local Similarity 79.2%; Pred No. 1.3e-224;
Matches 958; Conservative 0; Mismatches 237; Indels 15; Gaps 1;

```

OY 17 tcaagaaatggaagaacttgaagaatgtctctctgattccac-----gatcct 61
DB 4 tcaaggatggaagaacttcaagaatgtctctctgattccacgaagaactagaagttcca 63
OY 62 ctcatgaagtaacttgaacagacagagaagatcttggccaccctgtgtggaaccacagcg 121
DB 64 ttcacagaacacactgaaacagcacccagagatattgtcctctctctgtcggaacctgcgcg 123
OY 122 agtgaaccttccctccgtgtctgtgacctatgacctatctctctgtgtggggtaatg 181
DB 124 agccaatcttccctcccggtgtctgtgtgtatgtgtacatttgtgtgtgggggtacatt 183
OY 182 ggcacatctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 241
DB 184 ggcacatctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 243
OY 242 tactatctctcagcttggcagctcagatctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 301
DB 244 tactatctctcagcttggcagctcagatctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 303
OY 302 gaaatctagaagaatgtggaacaaattacaccttctctctgtgtgtgtgtgtgtgtgtgtgt 361
DB 304 gaggctcctagaagaatgtggaacaaattacaccttctctctgtgtgtgtgtgtgtgtgtgtgt 363
OY 362 aagacagccactctctcgaagactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 421
DB 364 aagacagccactctctcgaagactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 423
OY 422 gtagagcgtatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 481
DB 424 gtagagcgtatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 483
OY 482 cgggcccctcaggaatctctcagctcagctcagctcagctcagctcagctcagctcagctc 541
DB 484 cgggcccctcaggaatctctcagctcagctcagctcagctcagctcagctcagctcagctc 543
OY 542 accagacatcctcagctcagctcagctcagctcagctcagctcagctcagctcagctcagct 601
DB 544 accagacatcctcagctcagctcagctcagctcagctcagctcagctcagctcagctcagct 603
OY 602 gccacatcctcagctcagctcagctcagctcagctcagctcagctcagctcagctcagct 661
DB 604 gccacatcctcagctcagctcagctcagctcagctcagctcagctcagctcagctcagct 663
OY 662 ttcctcttctacatctctcctcctcagctcagctcagctcagctcagctcagctcagctc 721
DB 664 ttcctcttctacatctctcctcctcagctcagctcagctcagctcagctcagctcagctc 723
OY 722 aggtcgtgaagagaatgaatcccttgaagcgaacaaatgtgtgtgtgtgtgtgtgtgtgtgt 781
DB 724 aggtcgtgaagagaatgaatcccttgaagcgaatgaagagaatgaatgaatgaatgaatga 783
OY 782 tctagaagaatcagctcagctcagctcagctcagctcagctcagctcagctcagctcagct 841
DB 784 tctagaagaatcagctcagctcagctcagctcagctcagctcagctcagctcagctcagct 843
OY 842 accccctcctcagctcagctcagctcagctcagctcagctcagctcagctcagctcagct 901
DB 844 gccccgtcctcagctcagctcagctcagctcagctcagctcagctcagctcagctcagct 903
```

```

OY 902 gctgtctgtgttcaaccttcaatcctcagctcagctcagctcagctcagctcagctcagct 961
DB 904 gctgtctgtgttcaaccttcaatcctcagctcagctcagctcagctcagctcagctcagct 963
OY 962 gtcacacccattatctataactctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1021
DB 964 gtcacacccattatctataactctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1023
OY 1022 gtcctccctcagctcagctcagctcagctcagctcagctcagctcagctcagctcagct 1081
DB 1024 atctctcttctcacaacaaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1083
OY 1082 aagatcatcttctcagacgaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1141
DB 1084 cggaaatcttctcagacgaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1143
OY 1142 ttcctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1201
DB 1144 ttcctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1203
OY 1202 gtagcactaa 1211
DB 1204 atgtcaagaa 1213
```

RESULT 11
AAAA6022
ID AAAA6022 standard; cDNA; 1248 BP.

XX AAAA6022;
XX 22-AUG-2000 (first entry)

DE Human G protein coupled receptor hRup6 encoding cDNA SEQ ID NO:11.

XX Human; G protein coupled receptor; GPCR; transmembrane receptor;
KW Identification; agonist; screening; therapeutic; pharmaceutical;
XX mutant; ss.
XX OS Homo sapiens.

XX WO200022131-A2.

PD 20-APR-2000.

XX 13-OCT-1999; 99WO-US24065.

XX 13-OCT-1998; 98US-0170496.

XX 12-NOV-1998; 98US-0108029.

PR 20-NOV-1998; 98US-0109213.

PR 27-NOV-1998; 98US-0110060.

PR 16-FEB-1999; 99US-0120416.

PR 26-FEB-1999; 99US-0121852.

PR 12-MAR-1999; 99US-0123944.

PR 12-MAR-1999; 99US-0123945.

PR 12-MAR-1999; 99US-0123946.

PR 12-MAR-1999; 99US-0123948.

PR 12-MAR-1999; 99US-0123949.

PR 12-MAR-1999; 99US-0123951.

PR 28-MAY-1999; 99US-0136436.

PR 28-MAY-1999; 99US-0136437.

PR 28-MAY-1999; 99US-0136439.

PR 28-MAY-1999; 99US-0137127.

PR 28-MAY-1999; 99US-0137131.

PR 30-JUN-1999; 99US-0137567.

PR 27-AUG-1999; 99US-0141448.

PR 03-SEP-1999; 99US-0151114.

PR 29-SEP-1999; 99US-0152524.

PR 29-SEP-1999; 99US-0156633.

PR 29-SEP-1999; 99US-0156634.

XX (AREN-) ARENA PHARM INC.

XX Behan DP, Lehmann-Brinsma K, Chalmers DR, Chen R, Dang HT;
 PI Gore M, Llaw CM, Lin I, Lowitz K, White C;
 XX WPI: 2000-317986/27.
 DR P-PSDB: AAB02830.
 XX
 PT Non-endogenous, human G protein-coupled receptors for screening
 PT receptor, inverse or partial agonists useful as therapeutic agents -
 XX
 PS Example 1: Page 85-86; 187p; English.
 XX
 CC The present invention describes transmembrane receptors, preferably
 CC human G protein coupled receptors (GPCR), for which the endogenous
 CC ligand is unknown (orphan GPCR receptors). More specifically, the present
 CC invention relates to non-endogenous, constitutively activated versions
 CC of a human GPCR. These non-endogenous human GPCRs can be useful for
 CC the direct identification of candidate compounds as receptors agonists,
 CC inverse agonists or partial agonists for use as pharmaceutical agents.
 CC AAA46017 to AAA46126 and AAB02825 to AAB02859 represent sequences used in
 CC the exemplification of the present invention.
 CC
 SO Sequence 1248 BP: 269 A: 393 C: 269 G: 317 T: 0 other:

Query Match 64.2%; Score 790.8; DB 21; Length 1248;
 Best Local Similarity 79.2%; Pred. No. 1.3e-224;
 Matches 958; Conservative 0; Mismatches 237; Indels 15; Gaps 1;

OY 17 tcagtaatgaggaagaactggaagaatgcttcctggtacac-----gactcct 61
 DB 4 tcaaggaaatggaagaactctcagatgcttcctggtacacagagaagaactagaagaca 63
 OY 62 ctcatgaagtgacttgaaagacagagagatgagccaccttggtgagaccagagc 121
 DB 64 ttccagaaacacacagacagacagagagatcagccttccttcgagaccctggcgc 123
 OY 122 agtgaacataccctccggtgctgctgagcctagatccttcctggtgaggtaatg 181
 DB 124 agccactcttcctcccggtgctgctgagcctagatccttcctggtgaggtaatg 183
 OY 182 ggcacatctctggtgctgagatgctgcagacatcagacacttgaaagaccaccaa 241
 DB 184 ggcacatgctggtgctgctggtgctgcagacagcagctatgaagacgcccacaa 243
 OY 242 taccatcttcagcttgagcagatcagctgctgctgctgctggtgaggtcctcg 301
 DB 244 taccatcttcagcctgagcagctgctgacccctgctgctgctgctggtgaggtcctcg 303
 OY 302 gaaatcagagatgctgcacacatacccttcctgctgagcctggtgaggtacatc 361
 DB 304 gaggctctatgagatgctgcagacacatacccttcctgctgagcctggtgaggtacatc 363
 OY 362 aagacagcccttcctgagacgtgctgctgctcctcattcctgagtcacacaggtatgc 421
 DB 364 aagacagcccttcctgagacgtgctgctgctcctcattcctgagtcacacaggtatgc 423
 OY 422 gtgagagcgtctggtgagcatgctcagcccttcagcagcagctgagagagcgaggga 481
 DB 424 gtgagagcgtctggtgagcatgctcagcagcagctgagcagcagcgagcgagc 483
 OY 482 cgggagccctcagagcctcagcagcctgagcagcctcctgctgcttcttcctgccaat 541
 DB 484 cgggagccctcagagcctcctgagcagcctcctgagcagcctccttccttcctgccaat 543
 OY 542 accagacatcagatgagcagatgctgagcagccttcacacaggtgctcctgagtcagca 601
 DB 544 accagacatcagatgagcagatgctgagcagccttcacacaggtgctcctgagtcagca 603
 OY 602 ggcacctgagcaggtac 661
 DB 604 ggcacctgagcaggtac 663

OY 662 ttccctctacatcctcccaatgacccctcagcagcctcctcctacacatgaggtc 721
 DB 664 ttccatctcctacatcctcccaatgacccctcagcagcctcctcctacacatgaggtc 723
 OY 722 aggtcgaagagagatgaaatcccttgaaggcgaacaggtgctgtaataataacagacc 781
 DB 724 agactaagaagaagaatcctcttgagcagatgaaaggaagaaataataataacagacc 783
 OY 782 tctaaagaatcagatcagcagatgctgttctgctgctgctgctgctgctgctgctgctg 841
 DB 784 tgcagaagaatcagatcagcagatgctgttctgctgctgctgctgctgctgctgctgctg 843
 OY 842 accccctcagatgagcagcctcctcctcagccttctgaggaagatggaacagatccctg 901
 DB 844 gccccgttccaatatgacagcagcctcctcctcagccttctgaggaagatggaacagatccctg 903
 OY 902 gctgctggttcaacatcagcagatgctgagatcaggtctcttctatcagctcagc 961
 DB 904 gctgctggttcaacatcagcagatgctgagatcaggtctcttctatcagctcagc 963
 OY 962 gtcacccattatataacatcctcctgctgagccttcgagccttcgagccttcgaaatgt 1021
 DB 964 gtcacccattatataacatcctcctgctgagccttcgagccttcgagccttcgaaatgt 1023
 OY 1022 gtctccctacatcgaatggtgacatcccgagcctgcgacagagagcctcagaccag 1081
 DB 1024 atctctcttcccaaaacaggtgaccccgagcctgcgacagagagagcctcagaccag 1083
 OY 1082 aagatcattctctgacagaaatgctgagcctgctgagcctgacagagagatgagccacag 1141
 DB 1084 cggaaacattctcctgacagaaatgctgagcctgctgagcctgacagagagatgagccacag 1143
 OY 1142 ttccctgctgagatcagcagatcagcagcagcagcagcagcagcagcagcagcagcag 1201
 DB 1144 ttccctgctgagatcagcagatcagcagcagcagcagcagcagcagcagcagcagcag 1203
 OY 1202 gtacacataa 1211
 DB 1204 atgtcaagaa 1213

RESULT 12
 AAH49526
 ID AAH49526 standard; DNA: 1248 BP.
 XX
 AC AAH49526;
 DT 21-SEP-2001 (first entry)
 XX
 DE Human GTP-binding protein-coupled receptor GPRV39 coding sequence.
 XX
 KW GTP-binding protein-coupled receptor; neuroprotective; immunomodulatory;
 KW muscular; urinary; circulatory; anorectic; gene therapy; human;
 OS guanosine triphosphate; G-protein; ds.
 OS Homo sapiens.
 XX
 PN WO200148189-A1.
 PD 05-JUL-2001.
 XX
 PE 28-DEC-2000; 2000WO-JP09409.
 XX
 PR 28-DEC-1999; 99JP-0375152.
 PR 31-MAR-2000; 2000JP-0101339.
 PR 23-MAY-2000; 2000JP-0155978.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Matsumoto S, Oda T, Saito Y, Morikawa N, Yoshida K, Suwa M;
 PI Sugiyama T;
 XX WPI: 2001-425663/45.

DR P-PSDB: AAC64297.

XX Family of guanosine triphosphate binding protein coupled receptors and
PT genes encoding them for treatment and prevention of diseases associated
PT with these receptors.

PS Claim 1: Page 105, 137pp; Japanese.

XX The present sequence is the coding sequence for a human guanosine
CC triphosphate (GTP)-binding protein coupled receptor. The receptor is
CC useful for the investigation, diagnosis, treatment and prevention of
CC diseases associated with GTP-binding protein-coupled receptors, including
CC neurological, circulatory, digestive system, immune system, muscle and
CC urinary system disorders. GTP-binding proteins are also known as
CC G-proteins.

XX Sequence 1248 BP; 269 A; 393 C; 269 G; 317 T; 0 other:

Query Match 64.2%; Score 790.8; DB 22; Length 1248;

Best Local Similarity 79.2%; Pred. No. 1.3e-224; Mismatches 237; Indels 15; Gaps 1;

Matches 958; Conservative 0; Mismatches 237; Indels 15; Gaps 1;

QY 17 tcagtaatgggaaacttgaataatgtctcctgataccac-----gactct 61
DB 4 tcaggagatgaaataacttcgaatgtctcctgataccacgagaactagaagatcca 63
QY 62 ctatgaatgacttgaacagacagagagagacttggccacctgtgtgagaccagcgc 121
DB 64 ttccagaaacacacttgaacagacagagagagacttggccacctgtgtgagaccagcgc 123
QY 122 agtgaacctatcccttcgagtgctgtgagcctatgcttccctgtgaggggtaatg 181
DB 124 agcgaactctctctcccggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 183
QY 182 ggaacatctctctgt 241
DB 184 ggaatgtctctgt 243
QY 242 tactatctctctgt 301
DB 244 tactatctctctgt 303
QY 302 gaatctacagagaatgt 361
DB 304 gaagctatagagatgt 363
QY 362 aagaacagccctctcgaagactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 421
DB 364 aagaacagccctctcgaagactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 423
QY 422 gtagagagctatgt 481
DB 424 gtagagagctatgt 483
QY 482 cgggacctcagagatctcgaagactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 541
DB 484 cgggacctcagagatctcgaagactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 543
QY 542 accagatccatggaatcgaagactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 601
DB 544 accagatccatggaatcgaagactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 603
QY 602 ggaacatcgaagatcgaagactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 661
DB 604 ggaacatcgaagatcgaagactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 663
QY 662 ttccatctctacatctcgaagactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 721
DB 664 ttccatctctacatctcgaagactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 723
QY 722 aggcctgaagagagaatgaatcccttgaagcgaacaagctgtgtgtgtgtgtgtgtgtgt 781

DB 724 agactaagaagaacacatctcttgaagcagatgaaggaatgcaaatatccaagaacc 783
QY 782 tctagaagaatcagtcacacagatgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 841
DB 784 tgcagaataatcagtcacacagatgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 843
QY 842 acccctcagatggaacagcgcctctcgaagactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 901
DB 844 gcccggttccacatggaacagcgcctctcgaagactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 903
QY 902 gctgt 961
DB 904 gctgt 963
QY 962 gtaaccccatatctatcaacctcgt 1021
DB 964 gtaaccccatatctatcaacctcgt 1023
QY 1022 gtcctccctacatcgaatggt 1081
DB 1024 atctctctcttccacacacagtgagcctcccgatgacacacagtggtgtgtgtgtgtgtgtgtgtgtgtgt 1083
QY 1082 aagatcatctcttgaacagatgt 1141
DB 1084 cgaacatctctcgtgaacagatgt 1143
QY 1142 ttccctgt 1201
DB 1144 ttccctgt 1203
QY 1202 gtaacataa 1211
DB 1204 atgtcaagaa 1213

RESULT 13
AAH43075
ID AAH43075 standard; DNA; 1245 BP.
XX
AC AAH43075;
XX
DT 15-OCT-2001 (first entry)
XX
DE Nucleotide sequence of a human TGR-1 protein.
XX
KW TGR-1; neuromedin U; hypertension; stress disease; ss.
XX
OS Homo sapiens.
XX
FH
FT CDS 1..1245
FT /tag= a "TGR-1"
FT /product= "no termination codon given"
FT /note= "no termination codon given"
PN WO200157524 -A1.
XX
PD 09-AUG-2001.
XX
XX 02-FEB-2001; 2001WO-IP00746.
XX
PF 04-FEB-2000; 2000JP-0032773.
PR 24-FEB-2000; 2000JP-0052252.
PR 30-MAR-2000; 2000JP-0097896.
PR 19-JUN-2000; 2000JP-0187536.
XX
XX (TAKE) TAKEDA CHEM IND LTD.
XX
PI Hinuma S, Shintani Y, Hosoya M, Fujii R, Moriya T, Matsui H;
PI Okubo S;
XX
XX WPI: 2001-488917/53.
DR P-PSDB: AAC63366.

XX Identifying predicted or actual structures of two or more members of a
 PT chemical or physical library by mass spectrometry comprising
 PT correlating molecular mass measurements of two or more members with a
 PT shared chemical history -
 XX
 PS Disclosure: Page 85-87; 95pp; Japanese.
 XX
 CC The present sequence encodes a human TGR-1 protein. The specification
 CC describes a method of screening a compound, which is capable of binding
 CC properties of neuromedin U to TGR-1. The method is useful for screening
 CC preventatives and remedies for hypertension, stress diseases, etc..
 CC TGR-1 antagonists are also useful for treating the same diseases.
 CC
 XX Sequence 1245 BP; 268 A; 394 C; 268 G; 315 T; 0 other;

Query Match 64.0%; Score 787.6; DB 22; Length 1245;
 Best Local Similarity 79.0%; Pred. No. 1,le-223;
 Matches 956; Conservative 0; Mismatches 239; Indels 15; Gaps 1;

```

Oy 17 tcaagtaatgggaaactggaaatgctctcgtgatacgcac-----gactcct 61
Db 4 tcaaggatggaataaacttcagaaatgctcctcgtgatacgcacagagaactagaatcca 63
Oy 62 ctcaatgaagtaacttgaaacagacagagagatcacttgcccaactgtgtgaccacagcgc 121
Db 64 ttccgaataaacacactgaaacagacagagagatcacttgcccaactgtgtgaccacagcgc 123
Oy 122 agtgaactacccctcgcgtgtcgtgtgacatacgctgatactccctcgtgtggtggttaag 181
Db 124 agccactctctcctcccgctgtcgtgtgtgtatgtgatacgaatttctgtgtggtggtcatt 183
Oy 182 ggcaactctcctggtgtgatagtgtgtgtgacatacgaacttgaataacacccacac 241
Db 184 ggcaatgctcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 243
Oy 242 tactatctctcaagcttgtagcagctcacaatcgtcgtgtgtgtgtgtgtgtgtgtgtgtgtgt 301
Db 244 tactatctctcaagcttgtagcagctcacaatcgtcgtgtgtgtgtgtgtgtgtgtgtgtgtgt 303
Oy 302 gaaactcagagatgtgacacaaatcaccctctcctgtgtgtgtgtgtgtgtgtgtgtgtgtgt 361
Db 304 gaggctcactgagatgtgacacaaatcaccctctcctgtgtgtgtgtgtgtgtgtgtgtgtgtgt 363
Oy 362 aagacagccctctcctgaaacgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 421
Db 364 aagacagccctctcctgaaacgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 423
Oy 422 gtagagcgctatgtgacatgttcacaccccttcgcagacacagctgagagacagcgcgca 481
Db 424 gtagagcgctatgtgacatgttcacaccccttcgcagacacagctgagagacagcgcgca 483
Oy 482 cgggacctcagatcctcagcctagctgtgagctctcctgtgtgtgtgtgtgtgtgtgtgtgtgt 541
Db 484 cgggacctcagatcctcagcctagctgtgagctctcctgtgtgtgtgtgtgtgtgtgtgtgtgt 543
Oy 542 accgagcactcagatcagataaagttcagaacattcccaacgggtctcctcgtaactgtgtca 601
Db 544 accgagcactcagatcagataaagttcagaacattcccaacgggtctcctcgtaactgtgtca 603
Oy 602 gcaacactcagacatcaacacacacacacacacacacacacacacacacacacacacacacac 661
Db 604 gcaacactcagacatcaacacacacacacacacacacacacacacacacacacacacacacac 663
Oy 662 ttctctctcactacatcctcccaatgacccctcagcgtctcctcactacacacacacacacac 721
Db 664 ttctctctcactacatcctcccaatgacccctcagcgtctcctcactacacacacacacacac 723
Oy 722 aggcctgaagagagatgaatccctctgagcgcaacaaagtgtgtgtgtgtgtgtgtgtgtgtgt 781
Db 724 agacataaagaagaacaaatcctctctgagcgacatgagaggaatgcaaatatccaaagaccc 783

```

```

Oy 782 tcaagaagtcagtcaccacagatgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 841
Db 784 tgcagaataatcagtcacacagatgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 843
Oy 842 acccctcactcagtcagcagcgtctctcctcagccttctgtgaaagatgacagatccctcgt 901
Db 844 gccccgtccacatgacagcagcctctcctcagccttctgtgaaagatgacagatccctcgt 903
Oy 902 gctgtgtgttcacatcactatcagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 961
Db 904 gctgtgtgttcacatcactatcagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 963
Oy 962 gtaacacccatatacactcctcgtcgtcgcgtcgcgcgtcgcgcgtcgcgcgtcgcgcgtcgcgc 1021
Db 964 gtaacacccatatacactcctcgtcgtcgcgtcgcgcgtcgcgcgtcgcgcgtcgcgcgtcgcgc 1023
Oy 1022 gctcctcctacatcgaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1081
Db 1024 atctctcttcacacacacagtgacacacacacacacacacacacacacacacacacacacac 1083
Oy 1082 aagatcactctctgtgacgaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1141
Db 1084 cggaaacactctcctgtgacgaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1143
Oy 1142 ttccctgtgacatcactacacacacacacacacacacacacacacacacacacacacacac 1201
Db 1144 ttccctgtgacatcactacacacacacacacacacacacacacacacacacacacacacac 1203
Oy 1202 gtaccataaa 1211
Db 1204 atgtcaagaa 1213

```

RESULT 14
 AAF80326
 ID AAF80326 standard; DNA; 1594 BP.

AAF80326;

29-JUN-2001 (first entry)

Splice variant of G-protein coupled receptor IGS4A long version cDNA.
 KW Human: G-protein coupled receptor; IGS4; IGS4A; IGS4B; schizophrenia;
 KW nervous system disorder; psychiatric disorder; parkinson's disease;
 KW episodic paroxysmal anxiety; phobia; migraine; epilepsy; bulimia; stroke;
 KW cardiovascular disease; heart failure; angina pectoris; obesity; emesis;
 KW motility disorder; myocardial infarction; hypertension; dyslipidemia;
 KW gastrointestinal disorder; inflammatory bowel disease; osteoporosis;
 KW inflammation; infection; pain; cancer; immune disorder; allergy; sepsis;
 KW gynecological disorder; ss.

OS Homo sapiens.

Key Location/Qualifiers

FT CDS 55..945

FT /tag= a

FT /product= "splice variant of G-protein coupled receptor

FT IGS4A long version"

PN MO200125269-A2.

PD 12-APR-2001.

PF 25-SEP-2000; 2000WO-EP09584.

PR 24-SEP-1999; 99EP-0203140.

PR 24-SEP-1999; 99NL-1013140.

PR 28-JUL-2000; 2000EP-0202683.

PR 31-JUL-2000; 2000US-0222047.

XX (SOLV) SOLVAY PHARM BV.

PI Deleersnyder W, Berger C, Loeken C, Nys G, Venema J:
XX WPI: 2001-273568/28.
DR P-PSDB: AAB67806.
XX
PT New G-protein coupled receptors and the polynucleotides encoding them,
PT useful for preventing, ameliorating or correcting nervous system
PT disorders, cardiovascular diseases, dyslipidemias, inflammations, pain
PT or cancers
XX
PS Example 1b: Page 90-92; 102pp; English.
XX
CC The present sequence encodes a splice variant of the long version of a
CC human G-protein coupled receptor designated IGS4A. IGS4 exists in two
CC polymorphic forms, IGS4A and IGS4B. The IGS4 receptors and IGS4
CC polynucleotides are useful for preventing, ameliorating or correcting
CC dysfunctions or diseases. These diseases include peripheral nervous
CC system, psychiatric and central nervous system disorders
CC (e.g. schizophrenia, episodic paroxysmal anxiety, phobia, Parkinson's
CC disease, migraine, epilepsy, bulimia or stroke), cardiovascular diseases
CC (e.g. heart failure, angina pectoris, myocardial infarction or
CC hypertension), dyslipidemias, obesity, emesis, gastrointestinal
CC disorders (e.g. inflammatory bowel disease or motility disorders),
CC osteoporosis, inflammations, infections (e.g. bacterial, fungal,
CC protozoan or viral), pain, cancers, immune disorders, allergies,
CC sepsis or gynecological disorders. Agonists or antagonists of IGS4 are
CC effective with regard to disorders of the nervous system, including the
CC central and peripheral nervous systems, disorders of the gastrointestinal
CC system, cardiovascular system, skeletal muscle, thyroid, lung or
CC genitourinary system, or immunological disease. The IGS4 polynucleotides
CC are useful as diagnostic reagents for detecting under-expression,
CC overexpression or altered expression of IGS4.
XX
SQ Sequence 1594 BP: 370 A: 473 C: 332 G: 419 T: 0 other:

Query Match 55.0%: Score 677.6; DB 22: Length 1594;
Best Local Similarity 74.8%: Pred. No. 6e-191;
Matches 914: Conservative 0: Mismatches 229; Indels 79; Gaps 2;

QY 5 tggatttaagctcgttaatggaaacttgaaatgcttcctcggtacc----- 55
DB 46 tggattttaaagctcaggatgaaataacttcagatgcttcctcggtaccagcagaa 105
QY 56 -----gacccctcagtaagtaacttgaaacagacagagatgactgcccactgtg 109
DB 106 ctgaagatcattctcagaaacacccgaaacagacagagagatgactgcttcctctgc 165
QY 110 ggaaccaagcagtgacatccctccggtgctgctgacctatgctgactctctg 169
DB 166 ggaacctcgagcagcactctctccctccggtgctgctgacctatgctgactctctg 225
QY 170 gtgggggtatggcactctctggtgctgctgctgctgctgctgctgctgctgctgctg 229
DB 226 gtgggggtatggcactctctggtgctgctgctgctgctgctgctgctgctgctgctg 285
QY 230 aacaccacaactactatctctcagctgagctgacatgctgctgctgctgctgctgctg 289
DB 286 acgcccacaactactactctctcagctgagctgacatgctgctgctgctgctgctgctg 345
QY 290 ggaatgctcctggaatactcagagatgctgacatgctgctgctgctgctgctgctgctg 349
DB 346 ggaatgctcctggaatgctcagatgctgacatgctgctgctgctgctgctgctgctgctg 405
QY 350 ggaatgctcctggaatgctcagatgctgacatgctgctgctgctgctgctgctgctgctg 409
DB 406 ggaatgctcctggaatgctcagatgctgacatgctgctgctgctgctgctgctgctgctg 465
QY 410 aacacggttagcgttagagcgtatgctgacatgctgacatgctgacatgctgacatgctgac 469
DB 466 aacacggttagcgttagagcgtatgctgacatgctgacatgctgacatgctgacatgctgac 525
QY 470 agcaacggttagcgttagagcgtatgctgacatgctgacatgctgacatgctgacatgctgac 529

DB 526 agcacccgagcgcggccctcagatcctcggcagctgctggtctccgtgctctc 585
QY 530 tctttggccaataacagatcctcagatcctcggcagctgctggtctccgtgctctc 589
DB 586 tctttggccaataacagatcctcagatcctcggcagctgctggtctccgtgctctc 645
QY 590 gtaacctgctcagcaccctgacatgctgacatgctgacatgctgacatgctgacatgctgac 649
DB 646 gtaacctgctcagcaccctgacatgctgacatgctgacatgctgacatgctgacatgctgac 705
QY 650 caagctaccagctcctccttctacatcctcctcctcctcctcctcctcctcctcctcctc 709
DB 706 cagctaccctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 765
QY 710 ctcaatgggctcagctgtagagagatgctgacatgctgacatgctgacatgctgacatgctgac 769
DB 766 ctcaatgggctcagctgtagagagatgctgacatgctgacatgctgacatgctgacatgctgac 825
QY 770 attcacaagcctcctcagaaagcagctgacatgctgacatgctgacatgctgacatgctgac 829
DB 826 attcacaagcctcctcagaaagcagctgacatgctgacatgctgacatgctgacatgctgac 865
QY 830 gccatctgctgacccctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 869
DB 866 -----ccttggagagatgg 881
QY 890 acaagctcctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 949
DB 882 agtgaatcctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 941
QY 950 ctgaatcctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1009
DB 942 ctgaatcctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1001
QY 1010 ttctgaatgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1069
DB 1002 ttctgaatgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1061
QY 1070 ccttcagacccagaaagatcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 1129
DB 1062 caactcgccagcagaaagatcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 1121
QY 1130 gcaagcccccagctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 1189
DB 1122 atagtgccccaattcccatgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1181
QY 1190 tgtgcagagaggtacataaa 1211
DB 1182 tctagtgaacagatgcaagaa 1203
RESULT 15
AAF80327
ID AAF80327 standard; DNA: 1594 BP.
XX
XX AAF80327:
DT 29-JUN-2001 (first entry)
XX
DE Splice variant of G-protein coupled receptor IGS4A short version cDNA.
XX
XX Human; G-protein coupled receptor; IGS4; IGS4A; IGS4B; schizophrenia;
KW nervous system disorder; psychiatric disorder; Parkinson's disease;
KW episodic paroxysmal anxiety; phobia; migraine; epilepsy; bulimia; stroke;
KW cardiovascular disease; heart failure; angina pectoris; obesity; emesis;
KW motility disorder; myocardial infarction; hypertension; dyslipidemia;
KW gastrointestinal disorder; inflammatory bowel disease; osteoporosis;
KW inflammation; infection; pain; cancer; immune disorder; allergy; sepsis;
KW gynecological disorder; ss.
XX
XX Homo sapiens.
OS

Key Location/Qualifiers
 CDS 64..945 /*tag= a
 FT /product= "splice variant of G-protein coupled receptor
 FT IGS4 short version"
 PN MO200125269-A2.
 PD 12-APR-2001.
 XX 25-SEP-2000; 2000MO-EP09584.
 XX 24-SEP-1999; 99EP-0203140.
 PR 24-SEP-1999; 99NL-1013140.
 PR 28-JUL-2000; 2000EP-0202683.
 PR 31-JUL-2000; 2000US-0222047.
 XX (SOLV) SOLVAY PHARM BV.
 XX Deleersnyder W, Berger C, Loeken C, Nys G, Venema J:
 DR WPI: 2001-273568/28.
 DR P-PSDB; AAB67807.
 XX
 PT New G-protein coupled receptors and the polynucleotides encoding them,
 PT useful for preventing, ameliorating or correcting nervous system
 PT disorders, cardiovascular diseases, dyslipidemias, inflammations, pain
 PT or cancers -
 XX
 PS Example 1b: Page 94-95; 102Pp: English.
 XX
 CC The present sequence encodes a splice variant of the short version of a
 CC human G-protein coupled receptor designated IGS4. IGS4 exists in two
 CC polymorphic forms, IGS4A and IGS4B. The IGS4 receptors and IGS4
 CC polynucleotides are useful for preventing, ameliorating or correcting
 CC dysfunction or diseases. These diseases include peripheral nervous
 CC system, psychiatric and central nervous system disorders
 CC (e.g. schizophrenia, epilepsy, bulimia or stroke), cardiovascular diseases
 CC (e.g. heart failure, angina pectoris, myocardial infarction or
 CC hypertension), dyslipidemias, obesity, emesis, gastrointestinal
 CC disorders (e.g. inflammatory bowel disease or motility disorders),
 CC osteoporosis, inflammations, infections (e.g. bacterial, fungal,
 CC protozoan or viral), pain, cancers, immune disorders, allergies,
 CC sepsis or gynecological disorders. Agonists or antagonists of IGS4 are
 CC effective with regard to disorders of the nervous system, including the
 CC central and peripheral nervous systems, disorders of the gastrointestinal
 CC system, cardiovascular system, skeletal muscle, thyroid, lung or
 CC genitourinary system, or immunological disease. The IGS4 polynucleotides
 CC are useful as diagnostic reagents for detecting under-expression,
 CC overexpression or altered expression of IGS4.
 CC
 XX Sequence 1594 BP; 370 A; 473 C; 332 G; 419 T; 0 other:
 SO

Query Match 55.0%; Score 677.6; DB 22; Length 1594;
 Best Local Similarity 74.8%; Pred. No. 6e-191;
 Matches 914; Conservative 0; Mismatches 229; Indels 79; Gaps 2;

QY 5 tgaatttaagctcgaatgggaaacttgaataatgcttcctggaatccac----- 55
 DB tgaatttaatgtcagagatggaataacttcagaatgtctcctgatactacagaga 105
 QY 56 -----ggtcctcatgaatgaacttgaacagacagaggaatcttggccacctgtc 109
 DB ctagaagaatccatccagagaacacactgaacacgacagagagatcttgcctccctc 165
 QY 110 ggaaccaagcagatgaaccttcacccctccggtgtctgtgacctatgagcttctcctg 169
 DB 166 ggaacctgcgcagcgaaccttctcctcccggtgtctgtgacctatgagcttctcctg 225
 QY 170 gtggggtaattggcaatctctgtgtgcatgtgcatgttccgacatagaactttgaag 229
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

DB 226 gtggggtaattgcaatgtctcgtgtgtcctgtgtatcttcgaacacagctatgaag 285
 QY 220 aacaccacaactactatctcttcagctgtgcaagcctgaagatctgtctgtctgtt 289
 DB 286 agccccaactactactcttcacgcgcgtgtcctgcgaacctccgtgtccgtctc 345
 QY 290 ggaatgcctctgaatactacagagaatgtgacaaattacaccttctctgttgcggcgt 349
 DB 346 ggaatgcctcctgaggtctatagatgtggtgcgaactaacaccttctgttgcggcgt 405
 QY 350 ggaatgcctcctgaacagacgacctcttcgaagactgtgtgtccttgcctccatctcag 409
 DB 406 ggtctgacttcaagaagcgacctcttcgaagcgtgtgttctgcctccatctccagc 465
 QY 410 accaaggttaagatgaagcgtatgtgccaattgcaaccttcccgagcaagctgag 469
 DB 466 accacgctcaagcgtgagcgtacgtgccaactccacacccgcttccgcgcaaatgcag 525
 QY 470 agcagcgagcagcgccctcagagatcctcagacctagctgtgagcttctgtgtctt 529
 DB 526 agcaccgcgcgcgcgcgcctccagagatcctcgcgcacatcgtctggtgtcttcgtcttc 585
 QY 530 tcttggccaatatacagaatccatgagcatcaagtlccagacacttcccaacggtctcc 589
 DB 586 tccctgcaccaacacagacatccatgacatcaagtlccacacttccccaatgtgccgt 645
 QY 590 gtaactgtctcagcagctgcagatccacaaacccatgtgtgtatataactgtatc 649
 DB 646 gtcccaagtlcggcacgcgtacgtgacccaagcccatgtgtatatacaacttcatc 705
 QY 650 caagctacagcttctctctcacaatccccaagacccctacagctacgctccctacac 709
 DB 706 cagtgacactctctctctatctactctctcccaagcagctgataagtlctctcactac 765
 QY 710 ctcatgtggtcgaagcgtgaagagatgaatcccttgaagcgaacaaagtgtgtgaat 769
 DB 766 ctcatgtgctcactagactaaagaacaaatctcttgaagcagatgaagagatgcaat 825
 QY 770 attcaagaacctcagaagaagcagtcaccagaagatgtgtgttctgttctctgtt 829
 DB 826 attcaagaacctcagaagaacaaatccatcaacagaatgtgtt----- 865
 QY 830 gcatctgtcgaacccctccatcgtgacgcgctcttctcagcttctggaagaagt 889
 DB 866 -----cttggtaggaagtgtg 881
 QY 890 acagagctccctgctgtgtgttcaacctcatcatgtgtatcaggtgtctcttctat 949
 DB 882 agtgaatccctgtgtgtgttcaacctgcgtlcatgtgtgtcaggtgtctctctac 941
 QY 950 ctgagctcgcgtgtaaacccaattctataacctccctgtgtcgcggtctccgcgcgc 1009
 DB 942 ctgaagctcagcgtlcaaacccaattctataacctcactgtgtcgcgcgtctccagagaa 1001
 QY 1010 ttcgaaatgtgttccctcactacgcgaatgtgtccatccccgcgcgtccagagaa 1069
 DB 1002 ttccgaatgtgtatctctcttccacaacagtggtgcaatcccaatgacacacagt 1061
 QY 1070 cctccagcccaagaatcatctcttgaacagaatgtcacctcgtgtgagctgaagaagt 1129
 DB 1062 ccacgtgcccaagaaactcttccgtgacagaatgtccacttgtgtgagctgaagaagt 1121
 QY 1130 gcaagcccccagttccctgtgtagatcatcatcaacaacaaacttccacacgcgc 1189
 DB 1122 ataggtccccaatcccatgtatgataccaatgacaactcccaatcccaacgcctc 1181
 QY 1190 tgtgcagagaaggtacataaa 1211
 DB 1182 tctagtgaacagatgtcaagaa 1203

Job time: 6395 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2002, 14:19:18 : Search time 43.26 Seconds
(without alignments)
1335.589 Million cell updates/sec

Title: US-09-609-146-25
Perfect score: 2076
Sequence: 1 MCKLENASWTHDPLMKYLYNS.....GOSSIHNTLTAPCAGEVP 395

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP.ATHEA:*
2: SP.BACTERIA:*
3: SP.FUNGI:*
4: SP.HUMAN:*
5: SP.INVERTEBRATE:*
6: SP.MAMMAL:*
7: SP.MHC:*
8: SP.ORGANELLE:*
9: SP.PHAGE:*
10: SP.PLANT:*
11: SP.RODENT:*
12: SP.VIRUS:*
13: SP.VERTEBRATE:*
14: SP.UNCLASSIFIED:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2076	100.0	395	11 Q9JIB1	Q9JIB1 rattus norv
2	2063	99.4	395	11 Q9ESQ4	Q9ESQ4 rattus norv
3	1650.5	79.5	412	4 Q9NRAS	Q9NRAS homo sapien
4	1650.5	79.5	415	4 Q9GZ04	Q9GZ04 homo sapien
5	952.5	45.9	403	4 Q43664	Q43664 homo sapien
6	952.5	45.9	426	4 Q9H889	Q9H889 homo sapien
7	921.5	44.4	402	11 Q9JIB2	Q9JIB2 rattus norv
8	916.5	44.1	412	11 Q9JIB5	Q9JIB5 rattus norv
9	840	40.5	405	11 Q55040	Q55040 mus musculu
10	564.5	27.2	428	5 Q9VFNA	Q9VFNA drosophila
11	546.5	26.3	595	5 Q9VFW6	Q9VFW6 drosophila
12	525.5	23.3	660	5 Q9VFW5	Q9VFW5 drosophila
13	495.5	23.9	418	5 Q17239	Q17239 caenorhabd1
14	476.5	22.0	363	13 Q93413	Q93413 spherooides
15	470	22.6	374	13 Q93412	Q93412 spherooides
16	457.5	22.0	378	5 Q18701	Q18701 caenorhabd1
17	401	19.3	401	13 Q9DDR0	Q9DDR0 xenopus lae
18	400.5	19.3	454	4 Q9H573	Q9H573 homo sapien
19	399.5	19.2	400	6 Q9MYW9	Q9MYW9 macaca mula

20	398	19.2	397	13 Q9DDR1	Q9DDR1 xenopus lae
21	394	19.0	398	4 Q9UDE7	Q9UDE7 homo sapien
22	390	18.8	504	5 Q9VX8	Q9VX8 drosophila
23	388.5	18.7	398	4 Q9UDE6	Q9UDE6 homo sapien
24	388	18.7	404	13 Q9DFB0	Q9DFB0 catostomus
25	387	18.6	393	11 Q9R1M0	Q9R1M0 mus musculu
26	387	18.6	401	11 Q9R1L9	Q9R1L9 mus musculu
27	387	18.6	438	11 Q9R0D1	Q9R0D1 mus musculu
28	387	18.6	444	11 Q9J1Y1	Q9J1Y1 mus musculu
29	384.5	18.5	452	11 Q9JKN0	Q9JKN0 mus musculu
30	381.5	18.4	394	5 Q9U721	Q9U721 drosophila
31	379	18.3	377	13 Q9BU14	Q9BU14 brachydanio
32	376.5	18.1	383	13 Q42324	Q42324 catostomus
33	371.5	17.9	513	13 Q9DFA9	Q9DFA9 catostomus
34	368.5	17.8	367	13 Q9PVG0	Q9PVG0 carassius a
35	368.5	17.8	367	13 Q9PVF9	Q9PVF9 carassius a
36	366	17.6	384	13 Q9BUH1	Q9BUH1 brachydanio
37	365	17.6	559	13 Q9JK40	Q9JK40 mus musculu
38	360.5	17.4	385	11 Q9JK40	Q9JK40 mus musculu
39	360	17.3	388	4 Q9UIY1	Q9UIY1 homo sapien
40	354.5	17.1	352	11 Q98820	Q98820 rattus norv
41	354.5	17.1	352	11 Q9R297	Q9R297 rattus norv
42	354.5	17.1	352	11 Q9QWV3	Q9QWV3 rattus norv
43	354	17.1	425	6 Q9M2U2	Q9M2U2 oryctolagus
44	354	17.1	429	6 Q9M2U3	Q9M2U3 oryctolagus
45	347.5	16.7	466	11 Q9WU25	Q9WU25 cavia porce

ALIGNMENTS

RESULT	ID	PRELIMINARY:	PRT:	395 AA.
Q9JIB1	Q9JIB1			
AC	Q9JIB1			
DT	01-OCT-2000 (TREMBLrel. 15, Created)			
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)			
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)			
DE	NEUROMEDIN U RECEPTOR 2.			
GN	NM02K.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=SPRAGUE-DAWLEY;			
RX	MEDLINE=20351041; PubMed=10894543;			
RA	Howard A.D., Wang R., Pong S.-S., Mellin T.N., Strack A., Guan X.-M.,			
RA	Zeng Z., Williams D.L., Feigheimer S.D., Nunes C.N., Murphy B.,			
RA	Stair J.N., Yu H., Jiang Q., Clements M.K., Tan C.P., McKee K.K.,			
RA	Hreniuk D.L., McDonald T.P., Lynch K.R., Evans J.F., Austin K.P.,			
RA	Caskey T., van der Ploeg L.H.T., Liu Q.;			
RT	"Identification of receptors for neuromedin U and its role in			
RT	feeding.";			
RL	Nature 406:70-74(2000).			
DR	EMBL: AF242875; AAF82756.1; -			
DR	InterPro: IPR000276; GPCR_Kinopsn.			
DR	Pfam: PF00001; 7tm_1.1.			
DR	PRINTS: PR00237; GPCR_RHODPSN.			
DR	PROSITE: PS00237; G_PROTEIN_REC_P1_1; UNKNOWN_1.			
DR	PROSITE: PS50262; G_PROTEIN_REC_P1_2; 1.			
KW	Receptor.			
SO	SEQUENCE 395 AA; 44722 MW; 01D3765B5D5355C0 CRC64;			

QY	1	MCKLENASWTHDPLMKYLYNSTEERYLAHLCGPKRSDLSLPSVAVALIFLVGWNLLVCM	60
DB	1	MCKLENASWTHDPLMKYLYNSTEERYLAHLCGPKRSDLSLPSVAVALIFLVGWNLLVCM	60
Query Match	100.0%	Score 2076; DB 11; Length 395;	
Best Local Similarity	100.0%;	Pred. No. 1.3e-151;	
Matches	395;	Conservative 0; Mismatches 0; Indels 0; Gaps 0;	

```

OY 61 VIVRHOTLKTPTNYLFSFLAVSDLLVLLGMPLEIYEMHNPFLPGVCYFKTALFET 120
    |||||||
DB 61 VIVRHOTLKTPTNYLFSFLAVSDLLVLLGMPLEIYEMHNPFLPGVCYFKTALFET 120
OY 121 VCFASLSTVTTVSERYVAIVHFPRAKLESTRRRLALILSLVMSFSVFSLPMTSINGIK 180
    |||||||
DB 121 VCFASLSTVTTVSERYVAIVHFPRAKLESTRRRLALILSLVMSFSVFSLPMTSINGIK 180
OY 181 FOHPNGSSVPGSATCTVTKPMVYNLIIOATSFLEFIILPMTLISVLYLMLGRKDES 240
    |||||||
DB 181 FOHPNGSSVPGSATCTVTKPMVYNLIIOATSFLEFIILPMTLISVLYLMLGRKDES 240
OY 241 LEANKAVANIHRRPSRKSVTKMLFVLVLFALICWTPFHVDRLFESFVEEMTESLAAVENLI 300
    |||||||
DB 241 LEANKAVANIHRRPSRKSVTKMLFVLVLFALICWTPFHVDRLFESFVEEMTESLAAVENLI 300
OY 301 HVGSGVFYLLSAVNPIIYLLSRFRRAFRNVSPCKKCHPRHRPOGPPAKIIFLTE 360
    |||||||
DB 301 HVGSGVFYLLSAVNPIIYLLSRFRRAFRNVSPCKKCHPRHRPOGPPAKIIFLTE 360
OY 361 CHLVELTEDAGPOPGOSSIHNTNLITAPCAGEVP 395
    |||||||
DB 361 CHLVELTEDAGPOPGOSSIHNTNLITAPCAGEVP 395

```

RESULT 2

```

O9ES04 PRELIMINARY: PRT: 395 AA.
AC O9ES04: 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, last sequence update)
DE 01-JUN-2001 (Tremblrel. 17, last annotation update)
GN G-PROTEIN-COUPLED RECEPTOR TGR-1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN (1)
RP SEQUENCE FROM N.A.
RX PubMed=10887190;
RA Hosoya M., Moriya T., Kawamata Y., Ohkubo S., Fujii R., Matsui H.,
RA Shintani Y., Fukusumi S., Habata Y., Hinuma S., Onda H., Nishimura O.,
RA Fujino M.;
RT "Identification and Functional Characterization of a Novel Subtype of
RT Neumomedin U Receptor.";
RL J. Biol. Chem. 275:29528-29532(2000).
DR EMBL: AB041229; BAB13722.1; -
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PR00237; GPCRHHODOPSN.
DR PROSITE: PS00237; G-PROTEIN_RECP_F1_1; UNKNOWN_1.
DR PROSITE: PS50262; G-PROTEIN_RECP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 395 AA: 44756 MW: EABD6D36E9C355DA CRC64:

```

Query Match 99.4%; Score 2063; DB 11; Length 395;
 Best Local Similarity 99.2%; Pred. No. 1.3e-150;
 Matches 392; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

OY 1 MGLLENASWTHIDPLMKYLNSTEEYLAHLGCPKRSDSLSPVSAVALIFLVGMGNLLVCM 60
    |||||||
DB 1 MGLLENASWTHIDPLMKYLNSTEEYLAHLGCPKRSDSLSPVSAVALIFLVGMGNLLVCM 60
OY 61 VIVRHOTLKTPTNYLFSFLAVSDLLVLLGMPLEIYEMHNPFLPGVCYFKTALFET 120
    |||||||
DB 61 VIVRHOTLKTPTNYLFSFLAVSDLLVLLGMPLEIYEMHNPFLPGVCYFKTALFET 120
OY 121 VCFASLSTVTTVSERYVAIVHFPRAKLESTRRRLALILSLVMSFSVFSLPMTSINGIK 180
    |||||||
DB 121 VCFASLSTVTTVSERYVAIVHFPRAKLESTRRRLALILSLVMSFSVFSLPMTSINGIK 180

```

```

OY 181 FOHPNGSSVPGSATCTVTKPMVYNLIIOATSFLEFIILPMTLISVLYLMLGRKDES 240
    |||||||
DB 181 FOHPNGSSVPGSATCTVTKPMVYNLIIOATSFLEFIILPMTLISVLYLMLGRKDES 240
OY 241 LEANKAVANIHRRPSRKSVTKMLFVLVLFALICWTPFHVDRLFESFVEEMTESLAAVENLI 300
    |||||||
DB 241 LEANKAVANIHRRPSRKSVTKMLFVLVLFALICWTPFHVDRLFESFVEEMTESLAAVENLI 300
OY 301 HVGSGVFYLLSAVNPIIYLLSRFRRAFRNVSPCKKCHPRHRPOGPPAKIIFLTE 360
    |||||||
DB 301 HVGSGVFYLLSAVNPIIYLLSRFRRAFRNVSPCKKCHPRHRPOGPPAKIIFLTE 360
OY 361 CHLVELTEDAGPOPGOSSIHNTNLITAPCAGEVP 395
    |||||||
DB 361 CHLVELTEDAGPOPGOSSIHNTNLITAPCAGEVP 395

```

RESULT 3

```

O9NR6 PRELIMINARY: PRT: 412 AA.
AC O9NR6: 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, last sequence update)
DE 01-JUN-2001 (Tremblrel. 17, last annotation update)
GN NEURONMEDIN U RECEPTOR 2 (G-PROTEIN-COUPLED RECEPTOR TGR-1).
GN NMUR2 OR TGR-1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=20351041; PubMed=10894543;
RA Howard A.D., Wang R., Pong S.-S., Mellin T.N., Strack A., Guan X.-M.,
RA Zeng Z., Williams D.L., Feighner S.D., Nunes C.N., Murphy B.,
RA Stair J.N., Yu H., Jiang Q., Clements M.K., Tan C.P., McKee K.K.,
RA Hreniuk D.L., McDonald T.P., Lynch K.R., Evans J.F., Austin C.P.,
RA Caskey T., van der Ploeg L.H.T., Liu Q.;
RT "Identification of receptors for neumomedin U and its role in
RT feeding.";
RL Nature 406:70-74(2000).
RN (2)
RP SEQUENCE FROM N.A.
RX PubMed=10887190;
RA Hosoya M., Moriya T., Kawamata Y., Ohkubo S., Fujii R., Matsui H.,
RA Shintani Y., Fukusumi S., Habata Y., Hinuma S., Onda H., Nishimura O.,
RA Fujino M.;
RT "Identification and Functional Characterization of a Novel Subtype of
RT Neumomedin U Receptor.";
RL J. Biol. Chem. 275:29528-29532(2000).
DR EMBL: AF242874; AAF82755.1; -
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PR00237; GPCRHHODOPSN.
DR PROSITE: PS00237; G-PROTEIN_RECP_F1_1; UNKNOWN_1.
DR PROSITE: PS50262; G-PROTEIN_RECP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 412 AA: 47450 MW: 33D6D14A841B63E CRC64:

```

Query Match 79.5%; Score 1650.5; DB 4; Length 412;
 Best Local Similarity 79.4%; Pred. No. 5.8e-119;
 Matches 312; Conservative 36; Mismatches 40; Indels 5; Gaps 1;

```

OY 1 MGLLENASWTHIDPLMKYLNSTEEYLAHLGCPKRSDSLSPVSAVALIFLVGMGNLLVCM 55
    |||||||
DB 1 MGLLENASWTHIDPLMKYLNSTEEYLAHLGCPKRSDSLSPVSAVALIFLVGMGNLLVCM 55
OY 61 VIVRHOTLKTPTNYLFSFLAVSDLLVLLGMPLEIYEMHNPFLPGVCYFKTALFET 120
    |||||||
DB 61 VIVRHOTLKTPTNYLFSFLAVSDLLVLLGMPLEIYEMHNPFLPGVCYFKTALFET 120

```

QY 116 ALFETVCFASILSVTVSVRYVAIYHPRAKLESTRRRALRILSLVMSVSVFSLPNTS 175
 DB 121 ALFETVCFASILSVTVSVRYVAIYHPRAKLESTRRRALRILSLVMSVSVFSLPNTS 180
 QY 176 IHGKIFQHPNGSSVPGSATCTVTKPRMYYNLIQATSFLEYILPMTLSVLYLMGLRL 235
 DB 181 IHGKIFHPNGSLVPGSATCTVTKPRMYYNLIQATSFLEYILPMTLSVLYLMGLRL 240
 QY 236 KRDESLANKVAVNIHRPSKSVTKMLFVLVFAICMPFHVDRLFFSEVEEMTESLAA 295
 DB 241 KKDKSLEADGNNANIQPRKSVKMLFVLVFAICMAFPHIDRLFFSEVEEMTESLAA 300
 QY 296 VFNLHVSGVFYFYLSSAVNPITYNLSRRFRAFRNVSPCTCKMCHRRRPOGPAOKI 355
 DB 301 VFNLHVSGVFYFYLSSAVNPITYNLSRRFRAFRNVSPCTCKMCHRRRPOGPAOKI 360
 QY 356 IFLTECHLVLTEDAGPOFPQSSIHNTNLTTA 388
 DB 361 IFLTECHLVLTEDAGPOFPQSSIHNTNLTTA 393

RESULT 4
 Q9GZ04 PRELIMINARY: PRT: 415 AA.
 AC Q9GZ04:

DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE NEUROMEDIN U RECEPTOR 2 (NEUROMEDIN U RECEPTOR-TYPE 2).
 GN NMUR2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Plimates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN RN
 RP SEQUENCE FROM N.A.
 RX MEDLINE-20490668; PubMed-10899166;
 RA Raddatz R., Wilson A.E., Artyushyn R., Bonini J.A., Borowsky B.,
 RA Botchu L.W., Zhou S., Kouranova E.V., Nagorny R., Guevara M.S.,
 RA Dal M., Lerman G.S., Vayssie P.J., Branchek T.A., Gerald C., Forray C.,
 RA Adham N.;
 RT "Identification and Characterization of Two Neuropeptide U Receptors
 RT Differentially Expressed in Peripheral Tissues and the Central Nervous
 RT System";
 RT J. Biol. Chem. 275:32452-32459(2000).
 RN RN
 RP SEQUENCE FROM N.A.
 RX PubMed-11010960;
 RA Shan L., Qiao X., Crona J.H., Behan J., Wang S., Laz T., Bayne M.,
 RA Gustafson E.L., Monema F.J., Jr., Hedrick J.A.;
 RT "Identification of a Novel Neuropeptide U Receptor Subtype Expressed in
 RT the Central Nervous System";
 RT J. Biol. Chem. 275:39482-39486(2000).
 RN RN
 RP SEQUENCE FROM N.A.
 RA Pang L., Wang S., Laz T., Hedrick J.A.;
 RL Submitted (AUG-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AF217363; AAC24794.1; -;
 DR EMBL: AF292402; AAC03064.1; -;
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR Pfam: PF00001; 7tm.1. 1.
 DR PRINTS: PR00237; GPCR_Rhodopsn.
 DR PROSITE: PS00237; G_PROTEIN_RECPEP_FL_1; UNKNOWN_1.
 DR PROSITE: PS50262; G_PROTEIN_RECPEP_FL_2; 1.
 KW Receptor.
 SQ SEQUENCE 415 AA; 47725 MW; C2BACD84B313390F CRC64;

Query Match 79.5%; Score 1650.5; DB 4; Length 415;
 Best Local Similarity 79.4%; Pred. No. 5.8e-119;
 Matches 312; Conservative 36; Mismatches 40; Indels 5; Gaps 1;

QY 1 MGKLENAWTH-----DPLKKYLNSTEEYLAHCGPRKSDLSIPVAYALIFLVGMGN 55
 DB 4 MEKLOANASWITQOOLKEDPFOKHLNSTEEYLAFLCGPRRSHPFLPVSVVYPIPIVVGICGN 63
 QY 56 LVLGVNIHQTLKTPNTYVLSLAASDLLVLLGMPLEIYEMHNYPLFGVGYCFKT 115
 DB 64 VLVLGVNIHQAKTPTNYVLSLAASDLLVLLGMPLEIYEMHNYPLFGVGYCFKT 123
 QY 116 ALFETVCFASILSVTVSVRYVAIYHPRAKLESTRRRALRILSLVMSVSVFSLPNTS 175
 DB 124 ALFETVCFASILSVTVSVRYVAIYHPRAKLESTRRRALRILSLVMSVSVFSLPNTS 183
 QY 176 IHGKIFQHPNGSSVPGSATCTVTKPRMYYNLIQATSFLEYILPMTLSVLYLMGLRL 235
 DB 184 IHGKIFHPNGSLVPGSATCTVTKPRMYYNLIQATSFLEYILPMTLSVLYLMGLRL 243
 QY 236 KRDESLANKVAVNIHRPSKSVTKMLFVLVFAICMPFHVDRLFFSEVEEMTESLAA 295
 DB 244 KKDKSLEADGNNANIQPRKSVKMLFVLVFAICMAFPHIDRLFFSEVEEMTESLAA 303
 QY 296 VFNLHVSGVFYFYLSSAVNPITYNLSRRFRAFRNVSPCTCKMCHRRRPOGPAOKI 355
 DB 304 VFNLHVSGVFYFYLSSAVNPITYNLSRRFRAFRNVSPCTCKMCHRRRPOGPAOKI 363
 QY 356 IFLTECHLVLTEDAGPOFPQSSIHNTNLTTA 388
 DB 364 IFLTECHLVLTEDAGPOFPQSSIHNTNLTTA 396

RESULT 5
 O43664 PRELIMINARY: PRT: 403 AA.
 AC O43664:

DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE ORPHAN G PROTEIN-COUPLED RECEPTOR.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Plimates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN RN
 RP SEQUENCE FROM N.A.
 RA Tan C.T., McKee K.K., Palyha O.C., Feigmer S.D., Hreniuk D.L.,
 RA Smith R.G., Howard A.D.;
 RL Submitted (JAN-1998) to the EMBL/Genbank/DBJ databases;
 DR EMBL: AF044601; AAC02680.1; -;
 DR EMBL: AF044600; AAC02680.1; JOINED.
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR Pfam: PF00001; 7tm.1. 1.
 DR PRINTS: PR00237; GPCR_Rhodopsn.
 DR PROSITE: PS00237; G_PROTEIN_RECPEP_FL_1; UNKNOWN_1.
 DR PROSITE: PS50262; G_PROTEIN_RECPEP_FL_2; 1.
 SQ SEQUENCE 403 AA; 44980 MW; CF8D5F1A961CDBD9 CRC64;

Query Match 45.9%; Score 952.5; DB 4; Length 403;
 Best Local Similarity 47.9%; Pred. No. 1.9e-85;
 Matches 183; Conservative 60; Mismatches 94; Indels 45; Gaps 5;

QY 18 LNSREYTL-AHLGPKRSDLSIPVAVYALIFLVGMGNLVCVNIHQTLKTPNTYVL 76
 DB 17 LNLTDENLRKLYGPODTLFPICATYLLIFVGVAGNGLTCLVIRHAKMPTNYVL 76
 QY 77 FSLAVSDLLVLLGMPLEIYEMHNYPLFGVGYCFKTALFETVCFASILSVTVSVR 136
 DB 77 FSLAVSDLLVLLGMPLEIYEMHNYPLFGVGYCFRTLLFEVYCLASVLTALSVR 136
 QY 137 YVAIYHPRAKLESTRRRALRILSLVMSVSVFSLPNTSHGKIFQHPNGSSVPGSATC 196
 DB 137 YVAIYHPRAKLESTRRRALRILSLVMSVSVFSLPNTSHGKIFQHPNGSSVPGSATC 196
 QY 197 TVTKPMVYMLIQAATSFLEYILPMTLSVLYLMGLRKRDESL---EAN----- 244

RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Mei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mettel B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mlshina N.V., Mobary C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of *Drosophila melanogaster*."
 RL Science 287:2185-2195(2000)
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 DR EMBL: AE003703: AAF55016.1; -
 DR FLYbase: FBgn00384201: CG9918.
 DR InterPro: IPR000276: GPCR_Rhodopsn.
 DR Pfam: PF00001: 7tm.1: 1
 DR PRINTS: PR00237: GPCR_Rhodopsn.
 DR PROSITE: PS00237: G_PROTEIN_RECEP_F1_1: 1.
 DR PROSITE: PS00262: G_PROTEIN_RECEP_F1_2: 1.
 DR G-protein coupled receptor; Glycoprotein; Transmembrane.
 KW SEQUENCE 428 AA; 47198 MW; 9DD4A1B93046BD7 CRC64;

Query Match 27.2%; Score 564.5; DB 5; Length 428;
 Best Local Similarity 32.2%; Pred. No. 1.2e-35;
 Matches 118; Conservative 78; Mismatches 104; Indels 67; Gaps 6;

25 LAHLGPKRSDLS--LPVSVAVALFLVGVGNLWCMVVRHOTLTKPTNYLFLSAVS 82
 6 MSHDLGPPDPLAIVIPVYVYSLFITGVGNISCTIVAKKRSMTATNYLFLSLAIS 65
 83 DLAVLLGPLEIYEMWVHPFLGPGVCYFKTALFETVCFASILSVTVSVRYVAIVH 142
 66 DFLLLSGVQDEYSYIMSKRPYVGEVIGICRGLLAETSANAVLTITATVERYIAICH 125
 143 PPRAKLESTRRRRLRLSLVWSFVSPLPNTSHIGIKFOHPFNGSSVPGSATCTYTKPM 202
 126 PFLGQAMSKLSRIITIVIMAVIAIPOAOFGI--EHY-----SCVEQCGIYRVI 177
 203 WYVNLITQATSPFLYILPMTLSLVLYMGLRLKRDSELSA----- 243
 178 VKHS--FOLSTFLFLAPMSIILVLLYLGVLHLYLGVGPASVARROOLKSVSPDTIL 235
 244 -----NKVAVNIHRPSKRSYTKMLFVLVLEFAI 271
 236 YRGGSGTAMSFNGGSGAGTACLMGSSGAGQSSVGRRLNHYGTRRLRLRLVAVVYCFLL 295
 272 CWTPEVVDRLFEFSV---EEMTESLAAEVNLIHVVSGCFEYSSAVNPITIVLLSRFR 327
 296 CAAPFAORLAIYAPARGAKLKDQHEFVYVTVYVSGVLYSTCINPLLYNIMSHKFR 355
 328 AARNVV 334
 356 EAKRAVL 362

RESULT 11
 ID 09VFW6 PRELIMINARY; PRT; 595 AA.
 AC 09VFW6;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE CG8795 PROTEIN.
 GN CG8795
 OS *Drosophila melanogaster* (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Nephroidea; Drosophilidae; *Drosophila*.
 OC NCBI_Taxid=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE-20196006; PubMed-10731132;
 RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amandides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Mortan J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agdanyani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
 RA Ballaw R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokva D., Botchan M.R., Bouck J., Brokstein P., Bottler P.,
 RA Borkova D., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Mei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mettel B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mlshina N.V., Mobary C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of *Drosophila melanogaster*."
 RL Science 287:2185-2195(2000).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 DR EMBL: AE003699: AAF54929.1; -
 DR FLYbase: FBgn0038139: CG8795.
 DR InterPro: IPR000923: Copper_blue1.
 DR InterPro: IPR000276: GPCR_Rhodopsn.
 DR Pfam: PF00001: 7tm.1: 1.
 DR PRINTS: PR00237: GPCR_Rhodopsn.
 DR PROSITE: PS00196: COPPER_BLUE; UNKNOWN.1.
 DR PROSITE: PS00237: G_PROTEIN_RECEP_F1_1: 1.
 DR PROSITE: PS00262: G_PROTEIN_RECEP_F1_2: 1.
 KW G-protein coupled receptor; Glycoprotein; Transmembrane.
 SO SEQUENCE 595 AA; 65091 MW; 3E8E183BADF5471E CRC64;

Query Match 26.3%; Score 546.5; DB 5; Length 595;

RP SEQUENCE FROM N.A.
RA Palaya O.C., Tan C., McKee K.K., Felgner S.D., Hreniuk D.L.,
RA Smith R.G., Howard A.D.;
RT "Orphan G protein-coupled receptor from teleost fish Spheroides
RT nephelus related to growth hormone secretagogue receptor."
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF082209; AAC33472.1;
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR InterPro: IPR000524; HTH_Gntrk.
DR Pfam: PF00001; 7tm_1; 1.
DR Pfam: PR00237; GPCRHOODPSN.
DR PROSITE: PS00237; G_PROTEIN_RECP_F1.1; UNKNOWN_1.
DR PROSITE: PS00262; G_PROTEIN_RECP_F2.1.
DR PROSITE: PS00043; HTH_GNTR_FAMILY; UNKNOWN_1.
SO SEQUENCE 374 AA; 42324 MW; 2CF9304EF004C7A16 CRC64;

Query Match	22.6%;	Score 470;	DB 13;	Length 374;
Best Local Similarity	33.1%;	Pred. No. 1.8e-28;		
Matches 118;	Conservative 73;	Mismatches 117;	Indels 48;	Gaps 10;

```

OY 2 GKLENAASHIDPLKMYKJNSTEEXYLAHKCGKRSDLP---VSVALLELVGNGNL 57
Db 6 GLSPNCMW-----EGSHNGTAGLELPPLNNYISITPLLAIVTVACTVLETVGVGNWA 56
OY 58 VCAWIVRHQIOTKPTPNYYLFSLAUSDLLVLLGPLEIYEMHNHNPFLGCPVCYKRTAL 117
Db 57 TLVAVSRDRDRTTNNYLLCSMA5DLEIFVC-MPDLDMRWRYRPMREGDALCKLEOFV 115
OY 118 FETVCFASILSVTVSEYERVAIVHPPRAKLESTRRRLALSLWSF5VFLSPTMSIH 177
Db 116 SESCYSYSLITCALSYERYALALCEPLRAKALVTKRRVRALLILLMTVSL5SGVPFVWY 175
OY 178 GIKQKH--FPRGSSVPSGA-----TCTVTK-----PMKVYLLTIQANSELF 216
Db 176 GVEKDSTIMEFPSSDLN5SSWPLEAVDTRRECRMTQYAVESGLMEAMWL-----SSVF 227
OY 217 YILPMTLISVLYYMGLR--LKRD5LEANKVAVNHNRP5K5VYTKLEVLVFAICWT 274
Db 228 FEMPEVCTLVLYGLIGRRLMRHRETTINSRA--YHDK5NBOIQKMLVYVVLAVFLCWL 285
OY 275 PEHADR-LF5FE5EWM5ESLAAYNNLHV5GVCFEY5L5SAVNPIYVLL5RRPRA 329
Db 286 PEHAGRTQ5ESLDA5P5L5L5E5C5L5VA5LE5Y5A5INPL5VNT5MK5K5GA 341

```

Search completed: April 22, 2002, 14:19:19
Job time: 337 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2002, 14:13:39 ; Search time 42.45 Seconds
(without alignments)
689.257 Million cell updates/sec

Title: US-09-609-146-25

Perfect score: 2076
Sequence: 1 MKLEKNSWHDPLMKYLS.....GQSSIHNTNLTAPCAGEVP 395

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_1101.*
1: /SID8/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID8/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID8/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID8/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID8/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID8/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID8/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID8/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID8/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID8/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID8/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID8/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID8/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID8/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID8/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID8/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID8/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID8/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID8/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID8/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID8/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID8/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2076	100.0	395	22	AAE03634
2	2063	99.4	395	22	AAE03637
3	1650.5	79.5	412	22	AAE03634
4	1650.5	79.5	412	22	AAE03637
5	1650.5	79.5	412	22	AAE03634
6	1650.5	79.5	412	22	AAE03637
7	1650.5	79.5	412	22	AAE03634
8	1650.5	79.5	412	22	AAE03637
9	1650.5	79.5	412	22	AAE03634
10	1644.5	79.2	412	22	AAE03637
11	1640.5	79.0	412	22	AAE03634

12	1640.5	79.0	415	22	AAE03634
13	1632.5	78.6	415	22	AAE03637
14	1117.5	53.8	293	22	AAE03634
15	1117.5	53.8	296	22	AAE03637
16	1033.5	49.8	242	22	AAE03634
17	1022.5	49.3	249	22	AAE03637
18	952.5	45.9	403	21	AAE03634
19	952.5	45.9	403	21	AAE03637
20	952.5	45.9	403	22	AAE03634
21	952.5	45.9	403	22	AAE03637
22	946.5	45.6	403	22	AAE03634
23	916.5	44.1	412	22	AAE03634
24	916.5	44.1	413	22	AAE03634
25	916.5	44.1	439	22	AAE03634
26	840	40.5	405	21	AAE03634
27	840	40.5	405	22	AAE03634
28	499	24.0	418	17	AAE03634
29	477.5	23.0	412	21	AAE03634
30	477.5	23.0	412	22	AAE03634
31	477.5	23.0	412	22	AAE03634
32	476.5	23.0	363	21	AAE03634
33	476.5	23.0	363	22	AAE03634
34	472.5	22.8	412	21	AAE03634
35	465.5	22.4	364	18	AAE03634
36	465.5	22.4	364	22	AAE03634
37	465	22.4	353	18	AAE03634
38	465	22.4	353	18	AAE03634
39	465	22.4	364	21	AAE03634
40	463.5	22.3	361	18	AAE03634
41	463.5	22.3	362	18	AAE03634
42	463.5	22.3	364	18	AAE03634
43	463.5	22.3	366	21	AAE03634
44	463.5	22.3	366	21	AAE03634
45	463.5	22.3	366	22	AAE03634

ALIGNMENTS

RESULT 1	AAE03634	standard; Protein; 395 AA.
XX	AAE03634:	
XX	07-AUG-2001 (first entry)	
DE	Rat G-protein coupled receptor, SNORF72.	
KW	Rat; G-protein coupled receptor, SNORF72; neuromedin u neuropeptide;	
KW	MMU; inflammation; arthritis; autoimmune disease; septicemia; psychotic;	
KW	mental retardation; transplant rejection; neurological disorder; anxiety;	
KW	respiratory disorder; depression; schizophrenia; dementia; obesity; pain;	
KW	gastrointestinal disorder; hypertension; hypotension; epilepsy; diabetes;	
KW	ischemia; stroke; cancer; sexual disorder; circadian disorder; anorexia;	
KW	dermatological; psoriasis; Parkinson's disease; nausea; bulimia; allergy;	
KW	Alzheimer's disease; AIDS; hormonal disorder; memory disorder; migraine;	
KW	cardiovascular disorder; renal disorder; bone disease; delirium; asthma;	
KW	Cushing's disease; dysmenorrhoea; antidiabetic; osteoporosis;	
KW	metabolic disorder; behavioural disorder; Addison's disease; dyskinesia;	
KW	translucifer; antidiabetic; antidiabetic.	
XX	Rattus sp.	
OS		
XX		
FH	key	Location/Qualifiers
FT	Domain	40..63
FT	Domain	/label= Transmembrane_domain
FT	Domain	75..95
FT	Domain	/label= Transmembrane_domain
FT	Domain	106..132
FT	Domain	/label= Transmembrane_domain
FT	Domain	156..175
FT	Domain	/label= Transmembrane_domain

Amino acid sequenc
Human neurotensin-
Splice variant of
Amino acid sequenc
Human NGPCRI5, Ho
Human G-protein-co
Human growth hormo
Human FM-3. Homo
Human G-protein co
Human mutant G pro
Rat FM-3. Rattus
Rat G-protein coup
Rat G-protein coup
Mouse growth hormo
Murine FM-3. Mus
Human neurotensin
Amino acid sequenc
Long form of motil
Amino acid sequenc
The puffer fish mo
Amino acid sequenc
Human G-protein co
Rat growth hormone
Rat growth hormone
Pig growth hormone
Swine growth hormo
A mouse growth hor
Human growth hormo
Human growth hormo
Rat growth hormone
Human G-protein-co
Human G-protein-co
Rat growth hormone

```

FT Domain 202..227
FT /Label= Transmembrane_domain
FT 259..278
FT /Label= Transmembrane_domain
FT 292..322
FT /Label= Transmembrane_domain
PN W0200144297-A1.
XX
XX
PD 21-JUN-2001.
XX
XX
PF 13-DEC-2000; 2000MO-US33787.
XX
XX
PR 17-DEC-1999; 99US-0466435.
PR 25-APR-2000; 2000US-0558099.
PR 30-JUN-2000; 2000US-0609146.
XX
XX
PA (SYNA-) SYNAPTIC PHARM CORP.
PI Boulmi JA, Lerman GS, Quan Y, Ogozalek K;
XX
XX
DR WPI: 2001-390240/41.
DR N-PSDB; AAD08024.
XX
XX
PT A purified mammalian SNORF62 or SNORF72 receptor protein for
PT identification of compounds to treat e.g. inflammation, arthritis,
PT autoimmune diseases, transplant rejection, AIDS, cancer -
XX
XX
PS Claim 15; Fig 15; 256pp; English.
XX
XX
CC The invention relates to human G-protein coupled receptors, SNORF62 and
CC SNORF72 and their corresponding cDNA molecules. SNORF62 and SNORF72
CC receptors are specific for neuropeptide U (NMU) neuropeptides, hence they
CC are also known as NMU receptors. The agonist and antagonist of NMU
CC receptors are useful for treating an abnormality in a subject that is
CC alleviated by decreasing or increasing the activity of NMU receptor.
CC The NMU receptor serves as a valuable tool for designing drugs which are
CC useful for treating various pathological conditions such as
CC inflammation, arthritis, autoimmune diseases, transplant rejection,
CC graft vs host disease, bacterial, fungal, protozoan and viral infections,
CC septicemia, AIDS, pain, psychotic and neurological disorders, including
CC anxiety, depression, schizophrenia, dementia, mental retardation, memory
CC loss, epilepsy, neuromotor disorders, respiratory disorders, asthma,
CC eating/body weight disorders including obesity, bulimia, diabetes,
CC anorexia, nausea, hypertension, hypotension, vascular and cardiovascular
CC disorders, ischemia, stroke, cancers, sexual disorders, circadian
CC disorders, renal disorders, bone diseases including osteoporosis, benign
CC prostatic hypertrophy, gastrointestinal disorders, nasal congestion,
CC dermatological disorders such as psoriasis, allergies, Parkinson's
CC disease, Alzheimer's disease, acute heart failure, angina disorders,
CC delirium and dyskinesias such as Huntington's disease. They can also be
CC used to regulate steroid hormone disorders, epinephrine release
CC disorders, electrolyte balance disorders, endocrine disorders, memory
CC disorders, somatosensory disorders, metabolic disorders, behavioural
CC disorders, drug addiction, migraine, Addison's disease, Cushing's
CC disease, prevent miscarriage, induce labour or to treat dysmenorrhea.
CC The present sequence is rat G-protein coupled receptor, SNORF72 related
CC to the invention.
XX
XX
SQ Sequence 395 AA:

```

```

Query Match 100.0%; Score 2076; DB 22; Length 395;
Best Local Similarity 100.0%; Pred. No. 4, 9e-227;
Matches 395; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MGLEENASWIDPLMKYLNSTEEYLAHLGPKRSDDLSPVSVAALFLVGVMGNLLVCM 60
DB 1 mgleenaswihdpkmkylnsteeylahlgpkrsdslspvsvalfflvgnmgnllvcm 60
QY 61 VIVRHQTLPNTNYFLSLAVSDLLVLLGMPLEIYEMHNPPLFEGPVCCYKFTALFET 120
DB 61 vlvrhqtlkptnyflslavsdllvllgmpleyemhnpplfegpvgykftalfet 120

```

```

QY 121 VCFASILSVTVSVRYVAIVHPRAKLESTRRRALRLSLWSESVFSLPNTSINGIK 180
DB 121 vcfasilsvtlsvryvalvhpfraklestrrralrlslwsefsvfslpntsingik 180
QY 181 FOHPNGSSVPGSATCTVTKRMVWYNLIQATSEFLYFLPMTLISLVLYLMLGRKDES 240
DB 181 fohpngssvpgsatctvtkrmvwnyliqatseflyflpmtlislvllylmlgrkdes 240
QY 241 LEANKVAVNIHRPSRKSVTKMLFVLVLPALCWTFPHVDRLFFSEFEMTESLAAVNLI 300
DB 241 leankvavnihprsksvtkmlfvlvlpalcwtfphvdrlffsefemeslaavnli 300
QY 301 HVSQVFYFLSSAVNPITVNLSSRRFAAFNNVSPCKMCHPRRQGPQAQKIFLTE 360
DB 301 hvsgvfylssavnpitvnlssrrfaafnnvspckmchprrqgpqaqkiflte 360
QY 361 CHLVELTEDAGPOPPGQSSIHNTNLTTAPCAGEVP 395
DB 361 chlveltedagppfgqssihntnlttapcagevp 395

```

RESULT 2

```

AAG63367
ID AAG63367 standard; Protein: 395 AA.
XX
AC AAG63367;
XX
DT 15-OCT-2001 (first entry)
XX
XX Amino acid sequence of a rat TGR-1 protein.
XX
XX TGR-1; neuropeptide U; hypertension; stress disease.
XX
XX Rattus sp.
XX
XX W0200157524-A1.
XX
XX 09-AUG-2001.
XX
XX 02-FEB-2001; 2001MO-JP00746.
XX
XX 04-FEB-2000; 2000JP-0032773.
XX
XX 24-FEB-2000; 2000JP-0052252.
XX
XX 30-MAR-2000; 2000JP-0097896.
XX
XX 19-JUN-2000; 2000JP-0187536.
XX
XX (TAKE ) TAKEDA CHEM IND LTD.
XX
XX Hinuma S, Shintani Y, Hosoya M, Fujii R, Moriya T, Matsui H;
PI Okubo S;
XX
DR WPI: 2001-488917/53.
DR N-PSDB; AAA43078.
XX
XX Identifying predicted or actual structures of two or more members of a
XX chemical or physical library by mass spectrometry comprising
XX correlating molecular mass measurements of two or more members with a
XX shared chemical history -
XX
XX Claim 1; Page 89-91; 95pp; Japanese.
XX
XX The present sequence represents a rat TGR-1 protein. The specification
XX describes a method of screening a compound, which is capable of binding
XX properties of neuropeptide U to TGR-1. The method is useful for screening
XX preventatives and remedies for hypertension, stress diseases, etc..
XX TGR-1 antagonists are also useful for treating the same diseases.
XX
XX
SQ Sequence 395 AA:

```

```

Query Match 99.4%; Score 2063; DB 22; Length 395;
Best Local Similarity 99.2%; Pred. No. 1, 5e-225;

```

Matches 392; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```
OY 1 MGKLENAWIMHDPKMYNSTEETVLAHICGPKRSDLSLPVSAVALIFLYVGMNLVCM 60
DB 1 mglknaawimdpkmylnsteevylahicgpkrsdlsipvsavayallflvygmnllycm 60
OY 61 VVRRHOTKTPNNYLFSLAVSDLLVLLGMPLEIYEMHNPFLFGPGVGYKTALFET 120
DB 61 vvrhotktptnnylflslavsdllvllgmpleyemhnpflfgpgvcyktalfet 120
OY 61 vlvrtqltkprnyllflslavsdllvllgmpleyemhnpflfgpgvcyktalfet 120
DB 61 vlvrtqltkprnyllflslavsdllvllgmpleyemhnpflfgpgvcyktalfet 120
OY 121 VCFASISLVTVYVRYVAIVHPRAKLESTRRALRLISLWSFSVFSLPNTSIHGK 180
DB 121 vcfasislvtyvryvaivhpraklestrrralrlislwfsfvfslpntsihgk 180
OY 121 vcfasislvtyvryvaivhpraklestrrralrlislwfsfvfslpntsihgk 180
DB 121 vcfasislvtyvryvaivhpraklestrrralrlislwfsfvfslpntsihgk 180
OY 181 FOHPNGSSVPSATCTYTKPMWVYNLIIOATSPLEIYILPMLISLVLYMLRLKRD 240
DB 181 fohpngssvpsatctytkpwmvynliioat spleiyilpmlislvlymlrlkrd 240
OY 181 fohpngssvpsatctytkpwmvynliioat spleiyilpmlislvlymlrlkrd 240
DB 181 fohpngssvpsatctytkpwmvynliioat spleiyilpmlislvlymlrlkrd 240
OY 241 LEANKVAVNIHRSKSVTKMLFVLVFAICWTPFHVDRLFPSEVEWTESLAVFNL 300
DB 241 leankvavnihrsksvtkmlfvlvfaicwtpfhvdrlfpsevewteslaavfnl 300
OY 241 leankvavnihrsksvtkmlfvlvfaicwtpfhvdrlfpsevewteslaavfnl 300
DB 241 leankvavnihrsksvtkmlfvlvfaicwtpfhvdrlfpsevewteslaavfnl 300
OY 301 HVSQVFFYLSAVNPITIIYNLISRRFRAFRNVSPCTCKMCHPRHPOGPPAQKII 360
DB 301 hvsqvf fyllsavnpitiiynlisrrfafrnvspctckmchprhpoqpqakii 360
OY 301 hvsqvf fyllsavnpitiiynlisrrfafrnvspctckmchprhpoqpqakii 360
DB 301 hvsqvf fyllsavnpitiiynlisrrfafrnvspctckmchprhpoqpqakii 360
OY 361 CHLVLETDAGPQPGOSSIHNTNLTTAPCAGEVP 395
DB 361 chlveltdagpqpqgossihntnl ttpacagevp 395
OY 361 chlveltdagpqpqgossihntnl ttpacagevp 395
DB 361 chlveltdagpqpqgossihntnl ttpacagevp 395
```

RESULT 3

AAB67803
ID AAB67803 standard: Protein: 412 AA.

AC AAB67803:

DT 29-JUN-2001 (first entry)

DE Amino acid sequence of G-protein coupled receptor IGSA short version.

Human; G-protein coupled receptor; IGSA; IGSA4; IGSA4B; schizophrenia;
nervous system disorder; psychiatric disorder; Parkinson's disease;
episodic paroxysmal anxiety; phobia; migraine; epilepsy; bulimia; stroke;
cardiovascular disease; heart failure; angina pectoris; obesity; emesis;
mortality disorder; myocardial infarction; hypertension; dyslipidemia;
gastrointestinal disorder; inflammatory bowel disease; osteoporosis;
inflammation; infection; pain; cancer; immune disorder; allergy; sepsis;
gynecological disorder.

OS Homo sapiens.

PM WO200125269-A2.

PD 12-APR-2001.

PE 25-SEP-2000; 2000WO-EP09584.

PR 24-SEP-1999; 99EP-0203140.

PR 24-SEP-1999; 99NL-1013140.

PR 28-JUL-2000; 2000EP-0202683.

PR 31-JUL-2000; 2000US-0222047.

PA (SOLV) SOLVAY PHARM BV.

PI Deleersnyder W, Berger C, Loeken C, Nys G, Venema J;

DR MPI; 2001-273568/28.

DR N-PSDB; AAF80323.

PT New G-protein coupled receptors and the polynucleotides encoding them,
useful for preventing, ameliorating or correcting nervous system
disorders, cardiovascular diseases, dyslipidemias, inflammations, pain

PT or cancers -
XX Claim 18; Page 81-82; 102pp; English.

The present sequence represents the short version of a human G-protein
coupled receptor designated IGSA4. IGSA4 exists in two polymorphic forms,
CC IGSA4 and IGSA4B. The IGSA4 receptors and IGSA4 polynucleotides are useful
for preventing, ameliorating or correcting dysfunctions or diseases.
These diseases include peripheral nervous system, psychiatric and central
nervous system disorders (e.g. schizophrenia, episodic paroxysmal
anxiety, phobia, Parkinson's disease, migraine, epilepsy, bulimia or
stroke), cardiovascular diseases (e.g. heart failure, angina pectoris,
myocardial infarction or hypertension), dyslipidemias, obesity, emesis,
gastrointestinal disorders (e.g. inflammatory bowel disease or mortality
disorders), osteoporosis, inflammations, infections (e.g. bacterial,
fungal, protozoan or viral), pain, cancers, immune disorders, allergies,
sepsis or gynecological disorders. Agonists or antagonists of IGSA4 are
effective with regard to disorders of the nervous system, including the
central and peripheral nervous systems, disorders of the gastrointestinal
system, cardiovascular system, skeletal muscle, thyroid, lung or
gastrointestinal system, or immunological disease. The IGSA4 polynucleotides
are useful as diagnostic reagents for detecting under-expression,
overexpression or altered expression of IGSA4.

Sequence 412 AA;

Query Match 79.5%; Score 1650.5; DB 22; Length 412;
Best Local Similarity 79.4%; Pred. No. 1.le-178;
Matches 312; Conservative 36; Mismatches 40; Indels 5; Gaps 1;

```
OY 1 MGKLENAWIMH-----DPLMKYNSTEETVLAHICGPKRSDLSLPVSAVALIFLYVGMN 55
DB 1 mglknaawimhpyqkiedpdkhlnsteevylahicgpkrsdlsipvsavayallflvygmnl 60
OY 56 LVCWVYVNIHROTLPNTNYLFSLAVSDLLVLLGMPLEIYEMHNPFLFGPGVGYKT 115
DB 61 vlvclvvlhgdqamkprtnnylflslavsdllvllgmpleyemhnpflfgpgvcykt 120
OY 116 ALFETVCFASISLVTVYVRYVAIVHPRAKLESTRRALRLISLWSFSVFSLPNTS 175
DB 121 alfetvcfasislvtyvryvaivhpraklestrrralrlislwfsfvfslpntsihgk 180
OY 176 HGKIFRHPNGSSVPSATCTYTKPMWVYNLIIOATSPLEIYILPMLISLVLYMLRL 235
DB 181 hngkfhyfngssvpsatctytkpwmvynliioat spleiyilpmlislvlymlrl 240
OY 236 KRDESLKANKVAVNIHRSKSVTKMLFVLVFAICWTPFHVDRLFPSEVEWTESLA 295
DB 241 krdeslkankvavnihrsksvtkmlfvlvfaicwtpfhvdrlfpsevewteslaa 300
OY 296 VFNLIHVSQVFFYLSAVNPITIIYNLISRRFRAFRNVSPCTCKMCHPRHPOGPP 355
DB 301 vfnlihvsvqvf fyllsavnpitiiynlisrrfafrnvspctckmchprhpoqpq 360
OY 356 IFLTECHLVLETDAGPQPGOSSIHNTNLTTA 388
DB 361 ifltechlvletdagpqpqgossihntnl ttpa 393
```

RESULT 4

AAV71296
ID AAV71296 standard: Protein: 415 AA.

AC AAV71296:

DT 02-NOV-2000 (first entry)

DE Human orphan G protein-coupled receptor hRUP6.

PT Human; orphan G protein-coupled receptor; GPCR; hRUP6; drug screening;
transmembrane receptor; signal cascade.

XX

```
OS Homo sapiens.
XX WO200031258-A2.
XX
XX 02-JUN-2000.
XX
XX 13-OCT-1999. 99WO-US23687.
XX
XX 20-NOV-1998. 98US-0109213.
XX 16-FEB-1999. 99US-0120416.
XX 26-FEB-1999. 99US-0121852.
XX 12-MAR-1999. 99US-0123946.
XX 12-MAR-1999. 99US-0123949.
XX 28-MAY-1999. 99US-0136436.
XX 28-MAY-1999. 99US-0136437.
XX 28-MAY-1999. 99US-0136439.
XX 28-MAY-1999. 99US-0136567.
XX 28-MAY-1999. 99US-0137127.
XX 28-MAY-1999. 99US-0137131.
XX 29-JUN-1999. 99US-0141448.
XX 29-SEP-1999. 99US-0156555.
XX 29-SEP-1999. 99US-0156633.
XX 29-SEP-1999. 99US-0156634.
XX 29-SEP-1999. 99US-0156653.
XX 01-OCT-1999. 99US-0157280.
XX 01-OCT-1999. 99US-0157281.
XX 01-OCT-1999. 99US-0157282.
XX 01-OCT-1999. 99US-0157293.
XX 01-OCT-1999. 99US-0157294.
XX 12-OCT-1999. 99US-0416760.
XX 12-OCT-1999. 99US-0417044.
XX
XX (AREN-) ARENA PHARM INC.
XX
XX Chen R., Dang HT, Llaw CW, Lin I.
XX
XX WPI: 2000-400068/34.
XX N-PSDB: AAD01123.
XX
XX Novel human orphan G protein-coupled receptors and the encoding cDNAs
XX for use in the identification of G protein-coupled receptor agonists -
XX
XX Claim 22: Page 57-59; 102pp; English.
XX
XX The present amino acid sequence is the hrup6, an endogenous human
XX orphan G protein-coupled receptor (GPCR). The full length hrup6 cDNA
XX was cloned by RT-PCR using human thymus cDNA as template.
XX The orphan GPCR of the invention, like all GPCRs has seven transmembrane
XX alpha helices with an extracellular N-terminus and an intracellular
XX C-terminus. However, no endogenous ligands has yet been identified for
XX the proteins of the invention. The orphan GPCRs may be used in the
XX identification of their endogenous ligands, and to screen potential GPCR
XX agonists and antagonists for use as pharmaceutical agents. The proteins
XX may also be used in the study of GPCR-mediated signalling cascades, and
XX to elucidate their precise role in normal and diseased human conditions.
XX Nucleic acid encoding human orphan GPCRs may be used for tissue
XX localisation expression analysis to provide information about their
XX function in healthy and pathological states.
XX
XX Sequence 415 AA:
SQ
Query Match 79.5%; Score 1650.5; DB 21; Length 415;
Best Local Similarity 79.4%; Pred. No. 1, le-178;
Matches 312; Conservative 36; Mismatches 40; Indels 5; Gaps 1;
QY 1 MGKLENAWTH-----DPLMKYLNSTEEYLAHLGPKRSDLSLPVSAVALFLVGMGN 55
DB 4 meklgnasviyqkledpfqkhnsteeylaficgprsrhfflpvsvyplvtvgvign 63
QY 56 LLYCMYIVRHQTLKPTPNYLYSLAVSDLLVLLGMPLELYEMWNNHYRFLFGVGCYFKT 115
DB 64 vlycvtlqhqamkrcptnyltslavsdlvlllmpleyemwnrypflfgvgsyfxk 123
```

```
QY 116 ALPEYCFASILSVTVSEVERVAIVHPRAKLESTRRLRILSLWSFSVVSPLNTS 175
DB 124 alftvcfasilsvtvsvervyaillhpfraqlgsctrralrlglwgsfvsfpntcs 183
QY 176 IHGIRFOHPNGSSVPGSATCTVTKRPMWYNLIQATSEFLYLPTLISVLYLGLRL 235
DB 184 ihgirkfhypngslvpgsatctvtkrpmwlynliqatsefilylpmctvlsvlylmalrl 243
QY 236 KRDESLKANKVAVNIHRPSKSVTKMLFVLYLVFAICWPFHVDRLFFSFVEWTESLAA 295
DB 244 krdksleadeqnanlqrcpksvtnkmlfylvlfaicwafbhldrlffsfveewseslaa 303
QY 296 VFNLIHVSGVFYLLSAVNPITYNLSRFRFAFNWVSPCTCKMCHPRHPOGPPAOKI 355
DB 304 vfnlvhvsqvfylssavnpiltnlstrfqaafgnvlsfkkqwhsqhdpqjppaqn 363
QY 356 IFLECHLVELTEDAGPQFGGSSIHNTNLTTA 388
DB 364 ifltechveltedlqpfqgssmshnptra 396
RESULT 5
AAB02830
ID AAB02830 standard; Protein: 415 AA.
XX
XX AAB02830;
AC
XX 22-AUG-2000 (first entry)
DT
XX
XX Human G protein coupled receptor hrup6 protein SEQ ID NO:12.
DE
XX
XX Human: G protein coupled receptor; GPCR; transmembrane receptor;
XX identification; agonist; screening; therapeutic; pharmaceutical;
XX mutant.
XX
XX Homo sapiens.
OS
XX
XX WO200022131-A2.
PN
XX
XX 20-APR-2000.
PD
XX
XX 13-OCT-1999. 99WO-US24065.
PF
XX
XX 13-OCT-1998. 98US-0170496.
PR 12-NOV-1998. 98US-0108029.
PR 20-NOV-1998. 98US-0109213.
PR 27-NOV-1998. 98US-0110060.
PR 16-FEB-1999. 99US-0120416.
PR 26-FEB-1999. 99US-0121852.
PR 12-MAR-1999. 99US-0123944.
PR 12-MAR-1999. 99US-0123945.
PR 12-MAR-1999. 99US-0123946.
PR 12-MAR-1999. 99US-0123948.
PR 12-MAR-1999. 99US-0123949.
PR 12-MAR-1999. 99US-0123951.
PR 28-MAY-1999. 99US-0136436.
PR 28-MAY-1999. 99US-0136437.
PR 28-MAY-1999. 99US-0136439.
PR 28-MAY-1999. 99US-0137127.
PR 28-MAY-1999. 99US-0137131.
PR 28-MAY-1999. 99US-0137567.
PR 30-JUN-1999. 99US-0141448.
PR 27-AUG-1999. 99US-0151114.
PR 03-SEP-1999. 99US-0152524.
PR 29-SEP-1999. 99US-0156633.
PR 29-SEP-1999. 99US-0156655.
PR 29-SEP-1999. 99US-0156634.
XX
XX (AREN-) ARENA PHARM INC.
PA
XX Behan DP, Lehtmann-Bruinsma K, Chalmers DT, Chen R, Dang HT;
PI Gore M, Llaw CW, Lin I, Lowitz K, White C;
```

XX WP1: 2000-317986/27.
DR N-PSDB: AAA46022.
XX
XX Non-endogenous, human G protein-coupled receptors for screening
PT receptor, inverse or partial agonists useful as therapeutic agents -
XX
XX Example 1: Page 86-88; 187pp; English.
XX
XX The present invention describes transmembrane receptors, preferably
CC human G protein coupled receptors (GPCR), for which the endogenous
CC ligand is unknown (orphan GPCR receptors). More specifically the present
CC invention relates to non-endogenous, constitutively activated versions
CC of a human GPCR. These non-endogenous human GPCRs can be useful for
CC the direct identification of candidate compounds as receptors agonists,
CC inverse agonists or partial agonists for use as pharmaceutical agents.
CC AAA46017 to AAA46126 and AAB02825 to AAB02859 represent sequences used in
CC the exemplification of the present invention.
XX
XX Sequence 415 AA:

Query Match	79.5%	Score	1650.5	DB 21	Length	415
Best Local Similarity	79.4%	Pred. No.	1.1e-178			
Matches 312; Conservative	36	Mismatches	40	Indels	5	Gaps 1

```
Oy      1 MGKLENASWIH-----DELMKYLNSTEEYLAHLCGPKRSDLSIPVSVAVALIFELVGMGN 55  
        |::|::::|    |::|::::|::::|::|   |||||   |:|::|::|  
Db      4 meklqnaswlygqkledpfqkhlnsteeylaflcgprrshfflpvsvvyvpilfvvgvlgn 63
```

```
Oy      56 LRVCMVIVRIQTKTPTNYLYEFLASVDLLVLGLMPLEYEMWINYPLEFGPVGCEYKT 115
          :|||::||:|||||:|||||:|||||:||||:|||||:|||||:|||||:|||||:|||||:
Db      64 vlvclvllqhqamkrcptnylfslasvdllvlllqmpleyemwinyplfgpvgcyyfkt 123
```

```

Oy 116 ALFEVVCFASLI5VTVMSVERVVAIVHDFRAKLESTRARALRIIL5WMSFSVFSUPTS 1755
      |||||||:|||||:|||||:|||||: || |||:|||||
Db 124 alfevcfaasllstcvsveryvalhpfraqlqstrrarlllylwgfsvlfsipnts 1833

```

```

QY 176 IHGKEQHFPNGSSVPGSATCTVTKPMWVYLNLIQATSEFYILEMTLISVLYYLMGLRL 2355
      ||||| : ||||| ||||| ||||| : ||| ||||| : ||||| : ||||| |||
Db 184 lhgkfhfyfngslvpgsatctvklpmwlynlfiqvtstflfyllpmtvisvlylmalr| 2433

```

```
OY      236 KRDSLEANKAVAVNIHRPSRKSVTKMLFVLYLVAICMPHFHVDRLEFSFYEEWTESTLAA 2955
```

```
DB      244 kkdtsleadeqnanlqrpctsrsvnkmflvlylvafalcwbfhdrlffsfweesela 3033
```

```

OY 296 VFNLHVSGVEFYLTSSAVNPITLYNLSRRFRAAFRNVSPTCKMCHPRHRQGPRAOKI 355
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 304 VFNLHVSVGVFFYLTSSAVNPITLYNLSIRFGAAGANVSSFKKWSQHQDQIPQGRN 363

```

QY	356	IFTECHLVELTEDAGPQFPGOSIHNLT	388
Pb	364	ifftechfveltedjroffocasmhshl	396

RESULT 6
AAG63353

ID AAC03353 standard; protein; 415 aa.
 XX
 AC AAC63353;
 XX

DT 15-OCT-2001 (first entry)
 XX
 XX
 DE Amino acid sequence of a human TGR-1 protein.
 XX

KM	TGR-1; neuromedin U; hypertension; stress disease.
XX	
OS	Homo sapiens.
XY	

PN	MO20015/524-A1.
XX	
PD	09-AUG-2001.
XX	

```

PE      02-FEB-2001; 2001MO-JP00746.
XX
PR      04-FEB-2000; 2000JP-0032773.
PR      24-FEB-2000; 2000JP-0052252.
PR      30-MAR-2000; 2000JP-0097896.
PR      19-JUN-2000; 2000JP-0187536.
XX
PA      (TAKE ) TAKEDA CHEM IND LTD.
XX
PI      Hinuma S, Shintani Y, Hosoya M, Fujii R, Moriya T, Matsui H;
PI      Okubo S;
XX
DR      MPI: 2001-468917/53.
DR      N-PSTDB: AAH43072.
XX
PT      Identifying predicted or actual structures of two or more members of a
PT      chemical or physical library by mass spectrometry comprising
PT      correlating molecular mass measurements of two or more members with a
PT      shared chemical history -
XX
PS      Claim 1; Page 77-79; 95pp; Japanese.
XX
CC      The present sequence represents a human TGR-1 protein. The specification
CC      describes a method of screening a compound, which is capable of binding
CC      properties of neuromedin U to TGR-1. The method is useful for screening
CC      preventative and remedies for hypertension, stress diseases, etc..
CC      TGR-1 antagonists are also useful for treating the same diseases.
XX
SQ      Sequence 415 AA;

Query Match          79.5%; Score 1650.5; DB 22; Length 415;
Best Local Similarity 79.4%; Pred. No. 1.1e-178;
Matches 312; Conservative 36; Mismatches 40; Indels 5; Gaps 1;

QY      1 MGKLENASMIH-----DPLMKLYNSTEERYLAHLGCPKRSDLSLPSVAVALIFLVGN 55
    I:::IIIIII:: II :IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII
DB      4 meklgnaswlygqkledpfqghlnsteeylaflcgprshfflpvsyvvyrlfvgvlg 63
    ::::::::::::::::::::

QY      56 LLVCWVYVRHQIQLKPTNYTLFSLAVSDDLVLVLLGMPLEIYEEMHNYPFLGPVGCTYKT 115
    ::::::::::::::::::::
DB      64 vlvctvlilqhgmklrptnyllfsldavsdllvlllmplevyemwnypflfgpvgyckft 123
    ::::::::::::::::::::

QY      116 ALFEIVNCASILSVTVTSVERVALIVHPFRALKESTRRARLIISLWSFSVFSLPTS 175
    ::::::::::::::::::::
DB      124 alfeivcfaasilstvtvseveryallhpfraklgstrralrlilgviwgsvflsfpnts 183
    ::::::::::::::::::::

QY      176 IHGIKFQHPNGSSVGSATCTVTGPKMWVMNLIIGATSFLFYILPMTLISVLYLMGLRL 235
    IIIIIII::IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII
DB      184 ihgikfhfyngslvpgsatctvclpmwlifligvtsflfyllpmtvisvlylmairl 243
    ::::::::::::::::::::

QY      236 KRDESLKANNAVNIHRRSRKSRYTGMLEVLVLFACITPPRVNDLRFSPVEHWESLAA 295
    I:::IIIIII:: II II II II II II II II II II II II II II II II II II II
DB      244 kkdksleadeegnaalqrdrcksrvnmltvlvylvicwapflldtlflsfveeweesslaa 303
    ::::::::::::::::::::

QY      296 VFNILLHVVSQGVFFYLSSAVNPRIYNMLLSRRFAAFARNVVPCKWCCHPHRPQCIPAOKI 355
    ::::::::::::::::::::
DB      304 vfnllhvvsqgvffylssavnprilynmllsrrffaafarnvvpcpkwcchphrpqcipaqrn 363
    ::::::::::::::::::::

QY      364 IFLTECHLVELTEDAGPOPGOSSIHNTLTGA 388
    IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII
DB      364 ifltechveltedigpqfpgqssmhnhlpca 396
    ::::::::::::::::::::

RESUT# 7
AAG64297 ID AAG64297 standard; Protein: 415 AA.
XX
AC      AAG64297;
XX
DT      21-SEP-2001 (first entry)
XX
DE      Human CTR-binding protein-associated receptor CTR-30

```


CC nervous system disorders (e.g. schizophrenia, episodic paroxysmal
 CC anxiety, phobia, Parkinson's disease, migraine, epilepsy, bulimia or
 CC stroke), cardiovascular diseases (e.g. heart failure, angina pectoris,
 CC myocardial infarction or hypertension), dyslipidemias, obesity, emesis,
 CC gastrointestinal disorders (e.g. inflammatory bowel disease or motility
 CC disorders), osteoporosis, inflammations, infections (e.g. bacterial,
 CC fungal, protozoan or viral), pain, cancers, immune disorders, allergies,
 CC sepsis or gynecological disorders. Agonists or antagonists of IGS4 are
 CC effective with regard to disorders of the nervous system, including the
 CC central and peripheral nervous systems, disorders of the gastrointestinal
 CC system, cardiovascular system, skeletal muscle, thyroid, lung or
 CC genitourinary system, or immunological disease. The IGS4 polynucleotides
 CC are useful as diagnostic reagents for detecting under-expression,
 CC overexpression or altered expression of IGS4.

XX Sequence 412 AA:

Query Match 79.0%; Score 1640.5; DB 22; Length 412;
 Best Local Similarity 79.1%; Pred. No. 1.5e-177;
 Matches 311; Conservative 35; Mismatches 42; Indels 5; Gaps 1;

OY 1 MCKLENASMIH-----DPLMKYLNSTEEYLAHLGCPKRSDSLPSVAVYALIFLVGN 55
 DB 1 meklqnaswlygqkledpfqkhlnteeylaflcgrtrshfflpvsvvyplfivgyvign 60
 OY 56 LLVCMVIRHQTLPRTYVYLFSLAVSDLVLLGMPLEIYEMHNNYFPLGPPVCYKRT 115
 DB 61 vlvclvllqhgqamkprtyylfslavsdllvllgmplevemwnyplfipvgcyktc 120
 OY 116 ALFETVCASILSVTVTSVERVYVAIVHPRAKLESTRRRALRILSLWSFSVFSLEPMTS 175
 DB 121 alfevcfcsilslstvtvseveryallhpfraqlgstrrralllglvngfvsflsfpnts 180
 OY 176 IHGIRFOHFPNGSSVPGSNTCTVTRPMNVNLIOTATSEFLIPLMTLSVLYLMGRL 235
 DB 181 ihgikfhypngslvpgsatctvtrpnmwlylflqvtsflfyllpmtvlsvlylmalr 240
 OY 236 KRDESLANKVAVNTHRSRSKSVTKMLFVLVVFALICWTPFVVDLRFPSFEEMTESLAA 295
 DB 241 kdkdsleedegnanlqrcrsvnmkltfylvfalcaphldidllfsfveewteslaa 300
 OY 296 VFNLIHVSGVFEYLSAVNPITVNLISRRRAAFRNVSPTCKWCHRRHROPQPAQKI 355
 DB 301 vfnlvhvsgvfyfysavnpilynlisrrfqaatgvnysfthkqwhsqhdpqldpqdn 360
 OY 356 IFUTTECHLVELTEDAGPPPOGSSIHNTNLTTA 388
 DB 361 ifltechveltedipgpfllcqsavhshlpca 393

RESULT 12

AAB67804 AAB67804 standard; Protein: 415 AA.

XX AAB67804:

XX 29-JUN-2001 (first entry)

XX Amino acid sequence of a G-protein coupled receptor IGS4B long version.
 DE Human: G-protein coupled receptor; IGS4; IGS4A; IGS4B; schizophrenia;
 KW nervous system disorder; psychiatric disorder; Parkinson's disease;
 KW episodic paroxysmal anxiety; phobia; migraine; epilepsy; bulimia; stroke;
 KW cardiovascular disease; heart failure; angina pectoris; obesity; emesis;
 KW motility disorder; myocardial infarction; hypertension; dyslipidemia;
 KW gastrointestinal disorder; inflammatory bowel disease; osteoporosis;
 KW inflammation; infection; pain; cancer; immune disorder; allergy; sepsis;
 KW gynecological disorder.

OS Homo sapiens.
 XX
 PN WO200125269-A2.

XX 12-APR-2001.
 PD
 XX
 XX 25-SEP-2000; 2000WO-EP09584.
 PF
 XX 24-SEP-1999; 99EP-0203140.
 PR 24-SEP-1999; 99NL-1013140.
 PR 28-JUL-2000; 2000EP-0202683.
 PR 31-JUL-2000; 2000US-0222047.
 XX
 PA (SOLV) SOLVAY PHARM BV.
 XX
 PI Deleersnijder W, Berger C, Loeken C, Nys G, Venema J;
 PT WPI: 2001-273568/28.
 DR N-PSDB: AAF80324.
 XX
 PT New G-protein coupled receptors and the polynucleotides encoding them,
 PT useful for preventing, ameliorating or correcting nervous system
 PT disorders, cardiovascular diseases, dyslipidemias, inflammations, pain
 PT or cancers -
 XX
 XX Claim 19; Page 85-86; 102pp; English.

CC The present sequence represents the long version of a human G-protein
 CC coupled receptor designated IGS4B. IGS4 exists in two polymorphic forms,
 CC IGS4A and IGS4B. The IGS4 receptors and IGS4 polynucleotides are useful
 CC for preventing, ameliorating or correcting nervous system, psychiatric and
 CC these diseases include peripheral nervous system, psychiatric and central
 CC nervous system disorders (e.g. schizophrenia, episodic paroxysmal
 CC anxiety, phobia, Parkinson's disease, migraine, epilepsy, bulimia or
 CC stroke), cardiovascular diseases (e.g. heart failure, angina pectoris,
 CC myocardial infarction or hypertension), dyslipidemias, obesity, emesis,
 CC gastrointestinal disorders (e.g. inflammatory bowel disease or motility
 CC disorders), osteoporosis, inflammations, infections (e.g. bacterial,
 CC fungal, protozoan or viral), pain, cancers, immune disorders, allergies,
 CC sepsis or gynecological disorders. Agonists or antagonists of IGS4 are
 CC effective with regard to disorders of the nervous system, including the
 CC central and peripheral nervous systems, disorders of the gastrointestinal
 CC system, cardiovascular system, skeletal muscle, thyroid, lung or
 CC genitourinary system, or immunological disease. The IGS4 polynucleotides
 CC are useful as diagnostic reagents for detecting under-expression,
 CC overexpression or altered expression of IGS4.

XX Sequence 415 AA:

Query Match 79.0%; Score 1640.5; DB 22; Length 415;
 Best Local Similarity 79.1%; Pred. No. 1.5e-177;
 Matches 311; Conservative 35; Mismatches 42; Indels 5; Gaps 1;

OY 1 MCKLENASMIH-----DPLMKYLNSTEEYLAHLGCPKRSDSLPSVAVYALIFLVGN 55
 DB 4 meklqnaswlygqkledpfqkhlnteeylaflcgrtrshfflpvsvvyplfivgyvign 63
 OY 56 LLVCMVIRHQTLPRTYVYLFSLAVSDLVLLGMPLEIYEMHNNYFPLGPPVCYKRT 115
 DB 64 vlvclvllqhgqamkprtyylfslavsdllvllgmplevemwnyplfipvgcyktc 123
 OY 116 ALFETVCASILSVTVTSVERVYVAIVHPRAKLESTRRRALRILSLWSFSVFSLEPMTS 175
 DB 124 alfevcfcsilslstvtvseveryallhpfraqlgstrrralllglvngfvsflsfpnts 183
 OY 176 IHGIRFOHFPNGSSVPGSNTCTVTRPMNVNLIOTATSEFLIPLMTLSVLYLMGRL 235
 DB 184 ihgikfhypngslvpgsatctvtrpnmwlylflqvtsflfyllpmtvlsvlylmalr 243
 OY 236 KRDESLANKVAVNTHRSRSKSVTKMLFVLVVFALICWTPFVVDLRFPSFEEMTESLAA 295
 DB 244 kdkdsleedegnanlqrcrsvnmkltfylvfalcaphldidllfsfveewteslaa 303
 OY 296 VFNLIHVSGVFEYLSAVNPITVNLISRRRAAFRNVSPTCKWCHRRHROPQPAQKI 355
 IFUTTECHLVELTEDAGPPPOGSSIHNTNLTTA 388
 IFTECHVELTEDIPGPFLLCQSAVHSHLP 393

CC sepsis or gynecological disorders. Agonists or antagonists of IGSA are CC effective with regard to disorders of the nervous system, including the CC central and peripheral nervous systems, disorders of the gastrointestinal CC system, cardiovascular system, skeletal muscle, thyroid, lung or CC gallblutary system, or immunological disease. The IGSA polynucleotides CC are useful as diagnostic reagents for detecting under-expression, CC overexpression or altered expression of IGSA.

SQ Sequence 293 AA;

Query Match	53.8%	Score	1117.5	DB	22	Length	293
Best Local Similarity	79.0%	Pred. No.	2.4e-118				
Matches	211	Conservative	25	Mismatches	26	Indels	5
						Gaps	1

Oy		1	MGLTMSMWH-----DPLMKYLNSTEEYLTAHLCGPKRSQSLSPVSAVALFTLVGMGN	55
Dd		1	meklqnaswlygqrleobfqkhlmssteeyalficqprtsrhflpvbsvvypllvaylgn	60
Oy		56	LIVCWIWIRHOTLKTPTNYLYLESIAVSDDLVLGLMBLEIETEMHNYPFLGEVGCEFKT	115
Dd		61	vlvclvlilghqamktcrlpnuzylfsalavdlivlllgmblveyemwnrypfilgpygcfxk	120
Oy		116	ALFEPVCASLTSTNTYSVEREYVAIVHPFAKKESTRRRRLRILSLMWSSSVFSLPNTS	175
Dd		121	aifefcvcaasltstlcvsvegeyalihpfaklqsctrrlralfglvwgsvfslspnts	180
Oy		176	IHGIKFOHFPGSSVPSCATCYVRPMKWYNLLIQATSFFELYILPMTLSIVLYLMGLRL	235
Dd		181	iHgIkfhypfgsslvpgsaectvlkpmpwynfllqvcsffelyllpmelvsylvlylmnlr	240
Oy		236	KRDSELENKAAYAVIHHRPSKSVTXKL	262
Dd		241	kksklseedegnanlifpcrtsksvkm	267

RESULT	15
AAB67806	
ID	AAB67806 standard; Protein; 296 AA
XX	
AC	AAB67806;

Splice variant of G-protein coupled receptor IGS4A long version

KM Human; G-protein coupled receptor; ICS4; ICS34A; ICS4B; schizophrenia;
 KM nervous system disorder; psychiatric disorder; Parkinson's disease;
 KM epileptic paroxysmal anxiety; phobia; migraine; epilepsy; bulimia; stroke;
 KM cardiovascular disease; heart failure; angina pectoris; obesity; emesis;
 KM mobility disorder; myocardial infarction; hypertension; osteoporosis;
 KM gastrointestinal disorder; inflammatory bowel disease; osteoporosis;
 KM inflammation; infection; pain; cancer; immune disorder; allergy; sepsis;
 KM gynecological disorder.

OS	Homo sapiens.
XX	
PN	W0200125269-A2.

PD	12-APR-2001.	
XX		
PF	25-SEP-2000; 2000MO-EP09584	
XX		
PR	24-SEP-1999; 99NEP-0203140.	
PR	24-SEP-1999; 99NL-1013140.	
PR	28-JUL-2000; 2000EP-0202683	
PR	31-JUL-2000; 2000US-0222047.	

(SOLV) SOLVAY PHARM BV.

PI Deleersnyder W, Berger C, Loeken C, Nys G, Venema J;
XX
DR WPI: 2001-273568/28.

DR N-PSDB; AAB80326.

PT New G-protein coupled receptors and the polynucleotides encoding them,
PT useful for preventing, ameliorating or correcting nervous system
PT disorders, cardiovascular diseases, dyslipidemias, inflammations, pain
PT or cancers -

PS Example 1b; Page 92-93; 102pp; English.

The present sequence represents a splice variant of the long version of human G-protein coupled receptor designated IG54A. IG54 exists in two polymorphic forms, IG54A and IG54B. The IG54 receptors and IG54 polynucleotides are useful for preventing, ameliorating or correcting dysfunctions or diseases. These diseases include peripheral nervous system, psychophrenia, episodic paroxysmal anxiety, phobia, Parkinson's disease, migraine, epilepsy, bulimia or stroke), cardiovascular diseases (e.g. heart failure, angina pectoris, myocardial infarction or hypertension), dyslipidemia, obesity, emesis, gastrointestinal disorders (e.g. inflammatory bowel disease or motility disorders), osteoporosis, inflammations, infections (e.g. bacterial, fungal, protozoan or viral), pain, cancers, immune disorders, allergies, sepsis or gynecological disorders. Agonists or antagonists of IG54 are effective with regard to disorders of the nervous system. Including the central and peripheral nervous systems, disorders of the gastrointestinal system, cardiovascular system, skeletal muscle, thyroid, lung or genitourinary system, or immunological disease. The IG54 polynucleotides are useful as diagnostic reagents for detecting under-expression, overexpression or altered expression of IG54.

SQ Sequence 296 AA;

Query Match	53.8%	Score 1117.5	DB 22	Length 296
Best Local Similarity	79.0%	Pred. No. 2.4e-118		
Matches 211, Conservative	25	Mismatches 26	Indels 5	Gaps 1

[illegible]

Search completed: April 22, 2002, 14:13:39
Job time: 52 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2002, 14:12:47 ; Search time 42.45 Seconds
(without alignments)
724.156 Million cell updates/sec

Title: US-09-609-146-4
Perfect score: 2185
Sequence: 1 MSGMEKLQNASWYQOKLED.....ALSEQMSRTYQSFHFNKT 415

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues
Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 10%
Listing first 45 summaries

Database :

A.Geneseq_1101:*

- 1: /SID8/gcgdata/geneseq/geneseq/AA1980.DAT:*
- 2: /SID8/gcgdata/geneseq/geneseq/AA1981.DAT:*
- 3: /SID8/gcgdata/geneseq/geneseq/AA1982.DAT:*
- 4: /SID8/gcgdata/geneseq/geneseq/AA1983.DAT:*
- 5: /SID8/gcgdata/geneseq/geneseq/AA1984.DAT:*
- 6: /SID8/gcgdata/geneseq/geneseq/AA1985.DAT:*
- 7: /SID8/gcgdata/geneseq/geneseq/AA1986.DAT:*
- 8: /SID8/gcgdata/geneseq/geneseq/AA1987.DAT:*
- 9: /SID8/gcgdata/geneseq/geneseq/AA1988.DAT:*
- 10: /SID8/gcgdata/geneseq/geneseq/AA1989.DAT:*
- 11: /SID8/gcgdata/geneseq/geneseq/AA1990.DAT:*
- 12: /SID8/gcgdata/geneseq/geneseq/AA1991.DAT:*
- 13: /SID8/gcgdata/geneseq/geneseq/AA1992.DAT:*
- 14: /SID8/gcgdata/geneseq/geneseq/AA1993.DAT:*
- 15: /SID8/gcgdata/geneseq/geneseq/AA1994.DAT:*
- 16: /SID8/gcgdata/geneseq/geneseq/AA1995.DAT:*
- 17: /SID8/gcgdata/geneseq/geneseq/AA1996.DAT:*
- 18: /SID8/gcgdata/geneseq/geneseq/AA1997.DAT:*
- 19: /SID8/gcgdata/geneseq/geneseq/AA1998.DAT:*
- 20: /SID8/gcgdata/geneseq/geneseq/AA1999.DAT:*
- 21: /SID8/gcgdata/geneseq/geneseq/AA2000.DAT:*
- 22: /SID8/gcgdata/geneseq/geneseq/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2185	100.0	415	21	AAV71296
2	2185	100.0	415	21	AA802830
3	2185	100.0	415	22	AA663353
4	2185	100.0	415	22	AA664297
5	2185	100.0	415	22	AAE03629
6	2185	100.0	415	22	AA67802
7	2170	99.3	412	22	AA67803
8	2166	99.1	415	22	AA663366
9	2162	98.9	415	22	AA667804
10	2148	98.3	415	21	AAV52922
11	2147	98.3	412	22	AA67805

12	1650.5	75.5	395	22	AAE03634	Rat G-protein coup
13	1643.5	75.2	395	22	AA663367	Amino acid sequenc
14	1409	64.5	296	22	AA667806	Splice variant of
15	1394	63.8	293	22	AA667807	Splice variant of
16	1252	57.3	242	22	AA668333	Amino acid sequenc
17	1251	57.3	249	22	AA680937	Human ncPCR15. Ho
18	942	43.1	403	21	AAV90638	Human G-protein-co
19	942	43.1	403	21	AAV90673	Human mutant G pro
20	942	43.1	403	21	AAV44642	Human growth hormo
21	942	43.1	403	22	AA699185	Human FM-3. Homo
22	942	43.1	426	22	AAE03628	Human G-protein cou
23	884.5	40.5	413	22	AAE03635	Rat G-protein coup
24	880.5	40.3	412	22	AA699199	Rat FM-3. Rattus
25	880.5	40.3	439	22	AAE03636	Rat G-protein coup
26	825.5	37.8	405	21	AAV44641	Mouse growth hormo
27	825.5	37.8	405	22	AA699186	Murine FM-3. Mus
28	476	21.8	412	21	AAV54145	Amino acid sequenc
29	476	21.8	412	22	AA662652	Long form of motil
30	476	21.8	412	22	AA668478	Amino acid sequenc
31	474	21.7	412	22	AA802854	Human G-protein co
32	471	21.6	418	17	AA698562	Human neurotensin
33	469	21.5	363	21	AAV54147	The purifier fish mo
34	469	21.5	363	22	AA668479	Amino acid sequenc
35	453.5	20.8	400	22	AA668477	Amino acid sequenc
36	448	20.5	353	18	AAW19608	Pig growth hormone
37	448	20.5	353	18	AAW19215	Swine growth hormo
38	447	20.5	366	21	AAV90666	Human mutant G pro
39	444	20.3	361	18	AAW19217	Human growth hormo
40	444	20.3	362	18	AAW19610	Human growth hormo
41	444	20.3	366	21	AAV90632	Human G-protein-co
42	444	20.3	366	21	AAV70345	Human G-protein-co
43	444	20.3	366	22	AA697376	Rat growth hormone
44	444	20.3	366	22	AA662850	Human G-protein co
45	436.5	20.0	364	18	AAW19220	Rat growth hormone

ALIGNMENTS

RESULT 1	
ID	AAV71296 standard; Protein; 415 AA.
XX	
AC	AAV71296:
XX	
DT	02-NOV-2000 (first entry)
XX	
DE	Human orphan G protein-coupled receptor hRUP6.
XX	
KW	Human: orphan G protein-coupled receptor; GPCR; hRUP6; drug screening;
KW	Transmembrane receptor; signal cascade.
XX	
OS	Homo sapiens.
XX	
PN	WO200031258-A2.
XX	
PD	02-JUN-2000.
XX	
PF	13-OCT-1999; 99WO-US23687.
XX	
PR	20-NOV-1998; 98US-0109213.
PR	16-FEB-1999; 99US-0120416.
PR	26-FEB-1999; 99US-0121852.
PR	12-MAR-1999; 99US-0123946.
PR	12-MAR-1999; 99US-0123946.
PR	28-MAY-1999; 99US-0136436.
PR	28-MAY-1999; 99US-0136437.
PR	28-MAY-1999; 99US-0136437.
PR	28-MAY-1999; 99US-0136439.
PR	28-MAY-1999; 99US-0136567.
PR	28-MAY-1999; 99US-0137127.
PR	28-MAY-1999; 99US-0137131.
PR	29-JUN-1999; 99US-0141448.
PR	29-SEP-1999; 99US-0156555.

PR 29-SEP-1999; 99US-0156633.
 PR 29-SEP-1999; 99US-0156634.
 PR 29-SEP-1999; 99US-0156653.
 PR 01-OCT-1999; 99US-0157280.
 PR 01-OCT-1999; 99US-0157281.
 PR 01-OCT-1999; 99US-0157282.
 PR 01-OCT-1999; 99US-0157293.
 PR 01-OCT-1999; 99US-0157294.
 PR 12-OCT-1999; 99US-0416760.
 PR 12-OCT-1999; 99US-0417044.
 XX
 PA (AREN-) ARENA PHARM INC.
 XX
 PI Chen R, Dang HT, Llaw CM, Lin I;
 XX
 DR WPI: 2000-400068/34.
 DR N-PSDB: AAD01123.
 XX
 PT Novel human orphan G protein-coupled receptors and the encoding cDNAs
 for use in the identification of G protein-coupled receptor agonists -
 XX
 PS Claim 22: Page 57-59; 102pp: English.
 XX
 CC The present amino acid sequence is the hrup6, an endogenous human
 CC orphan G protein-coupled receptor (GPCR). The full length hrup6 cDNA
 CC was cloned by RT-PCR using human thymus cDNA as template.
 CC The orphan GPCR of the invention, like all GPCRs has seven transmembrane
 CC alpha helices with an extracellular N-terminus and an intracellular
 CC C-terminus. However, no endogenous ligands have yet been identified for
 CC the proteins of the invention. The orphan GPCRs may be used in the
 CC identification of their endogenous ligands, and to screen potential GPCR
 CC agonists and antagonists for use as pharmaceutical agents. The proteins
 CC may also be used in the study of GPCR-mediated signaling cascades, and
 CC to elucidate their precise role in normal and diseased human conditions.
 CC Nucleic acid encoding human orphan GPCRs may be used for tissue
 CC localisation expression analysis to provide information about their
 CC function in healthy and pathological states.
 CC
 XX
 SO Sequence 415 AA;
 Query Match 100.0%; Score 2185; DB 21; Length 415;
 Best Local Similarity 100.0%; Pred. No. 9, 4e-246;
 Matches 415; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 MSGMEKLNQNSMTYQOKLEDPFOKHLNSTEETYLAFICGPRSRHPELVSVYVPIFVGV 60
 Db 1 msgmeklqnaswlyqkledpfqkhlNSTEETYLAFICGPRSRHPELVSVVPIFVGV 60
 Oy 61 IGVNLVCLVILQHQAMKTPNNYLLFSLAVSDLVLLGMPLLEYEMWRNYPFLGPGVGY 120
 Db 61 IGVNLVCLVILQHQAMKTPNNYLLFSLAVSDLVLLGMPLLEYEMWRNYPFLGPGVGY 120
 Oy 121 FKTALEETVCFASITSTYTVSVRYVALILHFRKLOSTRRLRIIGIYMGSVLESLP 180
 Db 121 fktaleetvctfaststtytvsryvalilhfprklostrrlriigiymgsvlelslp 180
 Oy 181 NTSIHGKIFHYFPGSLVPSATCTVIKPMWYNIQVTSFLYLLPMTVISLYLMA 240
 Db 181 ntsihgkifhyfpgslvpsatctvikpmwynyiqvtsflyllpmtvislylylma 240
 Oy 241 LRLKKDKSLDEAGNANIQRPCRKSVNKMFLVLLVFAICWAPFHIDRLEFSEVEMSES 300
 Db 241 lrlkkdksldeagnanlqpcrksvnkmflvllvfaicwapfhidrlfseveemes 300
 Oy 301 LAAVFNLYVHYVSGVFFYLSSAVNPITYNLLSRPQAAFOVNVISFKQMSQHPOLPPA 360
 Db 301 laavfnlyvhvsgvffylssaavnpiitynllsrpqaafovnvissfhqmsqhpdpipa 360
 Oy 361 ORNIFLECHFEVLETDIGPOPCQSMHNSHLPALSSQMSQSTVNTQSHFNKT 415
 Db 361 qrniflechfevletdigpocqsmhnsahlpalssqmsqstvtntqshfnkt 415

RESULT 2
 AAB02830
 ID AAB02830 standard; Protein; 415 AA.
 XX
 AC AAB02830;
 XX
 DT 22-AUG-2000 (first entry)
 XX
 DE Human G protein coupled receptor hrup6 protein SEQ ID NO:12.
 XX
 KW Human; G protein coupled receptor; GPCR; transmembrane receptor;
 KW Identification; agonist; screening; therapeutic; pharmaceutical;
 KW mutant.
 XX
 OS Homo sapiens.
 XX
 PN WO200022131-A2.
 XX
 PD 20-APR-2000.
 XX
 PF 13-OCT-1999; 99MO-US24065.
 XX
 PR 13-OCT-1998; 98US-0170496.
 PR 12-NOV-1998; 98US-0108029.
 PR 20-NOV-1998; 98US-0109213.
 PR 27-NOV-1998; 98US-0110060.
 PR 16-FEB-1999; 99US-0120416.
 PR 26-FEB-1999; 99US-0120416.
 PR 12-MAR-1999; 99US-0123944.
 PR 12-MAR-1999; 99US-0123945.
 PR 12-MAR-1999; 99US-0123946.
 PR 12-MAR-1999; 99US-0123948.
 PR 12-MAR-1999; 99US-0123949.
 PR 12-MAR-1999; 99US-0123951.
 PR 28-MAY-1999; 99US-0136436.
 PR 28-MAY-1999; 99US-0136437.
 PR 28-MAY-1999; 99US-0136439.
 PR 28-MAY-1999; 99US-0137127.
 PR 28-MAY-1999; 99US-0137131.
 PR 28-MAY-1999; 99US-0137567.
 PR 30-JUN-1999; 99US-0141448.
 PR 27-AUG-1999; 99US-0151114.
 PR 03-SEP-1999; 99US-0152524.
 PR 29-SEP-1999; 99US-0156633.
 PR 29-SEP-1999; 99US-0156653.
 PR 29-SEP-1999; 99US-0156655.
 PR 29-SEP-1999; 99US-0156634.
 XX
 PA (AREN-) ARENA PHARM INC.
 XX
 PI Behan DP, Lehmann-Brulnasa K, Chalmers DT, Chen R, Dang HT;
 Gore M, Llaw CM, Lin I, Lowitz K, White C;
 XX
 DR WPI: 2000-317986/27.
 DR N-PSDB: AAA46022.
 XX
 PT Non-endogenous, human G protein-coupled receptors for screening
 receptor, inverse or partial agonists useful as therapeutic agents -
 XX
 PS Example 1: Page 86-88; 187pp: English.
 XX
 CC The present invention describes transmembrane receptors, preferably
 CC human G protein coupled receptors (GPCR), for which the endogenous
 CC ligand is unknown (orphan GPCR receptors). More specifically the present
 CC invention relates to non-endogenous, constitutively activated versions
 CC of a human GPCR. These non-endogenous human GPCRs can be useful for
 CC the direct identification of candidate compounds as receptors agonists,
 CC inverse agonists or partial agonists for use as pharmaceutical agents.
 CC AAA46017 to AAA46126 and AAB02825 to AAB02859 represent sequences used in
 CC the exemplification of the present invention.
 XX
 SO Sequence 415 AA;

Query Match 100.0%; Score 2185; DB 21; Length 415;
 Best Local Similarity 100.0%; Pred. No. 9,4e-246;
 Matches 415; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MSGMEKLNQASNIYQOKLEDPQKHUNSTEELAFGLCGPRSHFELPVSVVYVPPIFVGV 60
 1 msqmeklqnaswlygqkledpqkhlnsteeylafcgprshfelpsvvvyvpifvgy 60

61 IGVNLVCLVLIQHQAKPTNTNYLFSLAVSDLLVLLGMPLEVEYEMRNYPFLFGPGCY 120
 61 ignvnlvclvliqhqakptntnylfslavsdllvllgmpleveyemrnyplfgpgcy 120

121 FKTALFETVCFAISLITTVASERYVALHPFRAKLQSTRRAIRILGIYWGFSVLSLP 180
 121 fktalfetvcfasliltvavervallhpfraklqstrrairilgiywgfsvlslp 180

181 NTSIGIKFHFPPNGSLVPGSATCTVIKPMWIIYNTSFLFYLLPMTVISVLYYLA 240
 181 ntsigikfhfppngslvpgsatctvikpmwiiyntsflyllpmvisvlyylma 240

241 LRLKDKSLADEGNANIQRCRKSVNKMFLVLYVFAICWAPPHIDRLFFSFEEMSES 300
 241 lrlkdksladegnanlqrcrksvnkmflvlyvfaicwapphidrlffsfeewses 300

301 LAAVFNLVHVSGVFYLLSSAVNPITYNLSRRFOAFQNVISFHKQMSQHDPLPPA 360
 301 laavfnlvhvsgvfylssavnpitynlsrrfoafqnvissfhkqmsqhdplppa 360

361 ORNIFLTECHVELTEDIGPOFQOSSMHNHSLPTALSSQMSRTNYQSDFHFKT 415
 361 qrnifltechveltedigpofqossmhnhslptalsseqmsrtnyqsfhfk 415

RESULT 3
 AAG63353 standard; Protein: 415 AA.
 ID AAG63353
 AC AAG63353:
 DT 15-OCT-2001 (first entry)
 DE Amino acid sequence of a human TGR-1 protein.
 KM TGR-1; neuromedin U; hypertension; stress disease.
 OS Homo sapiens.
 PN WO200157524-A1.
 PD 09-AUG-2001.
 PF 02-FEB-2001; 2001WO-JP00746.
 PR 04-FEB-2000; 2000JP-0032773.
 PR 24-FEB-2000; 2000JP-0052252.
 PR 30-MAR-2000; 2000JP-0097896.
 PR 19-JUN-2000; 2000JP-0187536.
 PA (TAKE) TAKEDA CHEM IND LTD.
 PI Hinuma S, Shintani Y, Hosoya M, Fujii R, Moriya T, Matsui H.
 PI Okubo S;
 DR MPI; 2001-488917/53.
 DR N-PSDB; AAH43072.
 XX Identifying predicted or actual structures of two or more members of a
 PT chemical or physical library by mass spectrometry compising
 PT correlating molecular mass measurements of two or more members with a
 PT shared chemical history -
 PS Claim 1; Page 77-79; 95pp; Japanese.

XX The present sequence represents a human TGR-1 protein. The specification
 CC describes a method of screening a compound, which is capable of binding
 CC properties of neuromedin U to TGR-1. The method is useful for screening
 CC preventatives and remedies for hypertension, stress diseases, etc..
 CC TGR-1 antagonists are also useful for treating the same diseases.
 XX Sequence 415 AA:

Query Match 100.0%; Score 2185; DB 22; Length 415;
 Best Local Similarity 100.0%; Pred. No. 9,4e-246;
 Matches 415; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MSGMEKLNQASNIYQOKLEDPQKHUNSTEELAFGLCGPRSHFELPVSVVYVPPIFVGV 60
 1 msqmeklqnaswlygqkledpqkhlnsteeylafcgprshfelpsvvvyvpifvgy 60

61 IGVNLVCLVLIQHQAKPTNTNYLFSLAVSDLLVLLGMPLEVEYEMRNYPFLFGPGCY 120
 61 ignvnlvclvliqhqakptntnylfslavsdllvllgmpleveyemrnyplfgpgcy 120

121 FKTALFETVCFAISLITTVASERYVALHPFRAKLQSTRRAIRILGIYWGFSVLSLP 180
 121 fktalfetvcfasliltvavervallhpfraklqstrrairilgiywgfsvlslp 180

181 NTSIGIKFHFPPNGSLVPGSATCTVIKPMWIIYNTSFLFYLLPMTVISVLYYLA 240
 181 ntsigikfhfppngslvpgsatctvikpmwiiyntsflyllpmvisvlyylma 240

241 LRLKDKSLADEGNANIQRCRKSVNKMFLVLYVFAICWAPPHIDRLFFSFEEMSES 300
 241 lrlkdksladegnanlqrcrksvnkmflvlyvfaicwapphidrlffsfeewses 300

301 LAAVFNLVHVSGVFYLLSSAVNPITYNLSRRFOAFQNVISFHKQMSQHDPLPPA 360
 301 laavfnlvhvsgvfylssavnpitynlsrrfoafqnvissfhkqmsqhdplppa 360

361 ORNIFLTECHVELTEDIGPOFQOSSMHNHSLPTALSSQMSRTNYQSDFHFKT 415
 361 qrnifltechveltedigpofqossmhnhslptalsseqmsrtnyqsfhfk 415

RESULT 4
 AAG64297 standard; Protein: 415 AA.
 ID AAG64297
 AC AAG64297:
 DT 21-SEP-2001 (first entry)
 DE Human GTP-binding protein-coupled receptor GPR39.
 KM GTP-binding protein-coupled receptor; neuroprotective; immunomodulatory;
 KM muscular; urinary; circulatory; anorectic; human; guanosine triphosphate;
 G-protein.
 OS Homo sapiens.
 PN WO200148189-A1.
 PD 05-JUL-2001.
 PF 28-DEC-2000; 2000WO-JP09409.
 PR 28-DEC-1999; 99JP-0375152.
 PR 31-MAR-2000; 2000JP-0101339.
 PR 23-MAY-2000; 2000JP-0155978.
 PA (HELI-) HELIX RES INST.
 PI Matsumoto S, Oda T, Satou Y, Morikawa N, Yoshida K, Suwa M;
 PI Sugiyama T;

Query Match	100.0%;	Score 2185;	DB 22;	Length 415;
Best Local Similarity	100.0%;	pred. NO. 9.4e-246;		
Matches 415; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0.

Qy	1	MSGNEKIQNANSMITQOAKLEDPEFOKHLNSTRYEALICGPRASHRELPVSYVYDIFPVWG	60
Db	1	msgnek1qnaswsiyqkledpfqkhlustryeaelicgprashre1pvsyvypdlfvwg	60
Qy	61	IGNVLCVTLVLOHQAOKTPTNXYLFSLAVSDLVLLCMLPELEVEMRNRPDLGPVGY	120
Db	61	IGNVLCVTLVlqhgaMktrpnyllfslavsdllvlllqmpleyeamwnryfllgprgy	120
Qy	121	FKTALFETVCFASLISITTVSYVERVYALHPRAKLOSTRRALRICITWGSVLESLP	180
Db	121	fktalfetvcfaslislsttvsveryaallhpfreaklqstrralrlilg1vngfavlslp	180
Qy	181	NTSHGIRKFNHPFNGSLVPBSATCYIKRMWYTNFIQVTSFLFLLPMYTIASVYYLMA	240
Db	181	ntshgirkfnhpfngslvpbsatcvcykrmwtnfnfiqvtsflfllpmytiavyyylma	240
Qy	241	LRIKKDSLEADEGNANIORPCRSVNMKLFUYLVLFALCAAPFHIDRLFFSFVEEMSES	300
Db	241	lrikkdsleadegnanhlqprckrsynkm1fvlvylfalcwprfhndrlffsfveeases	300
Qy	301	LAAYENLVHVVSQGVFFYLSSAVNDIYNLLSRFQAAFQNVTSRFXKQNHSHQHDQLEPPA	360
Db	301	laayenlvhvvsqgvffylsavnpdl1ynllstrfgaafqnvtsrfkqnhshdqhdppa	360
Qy	361	ORNLFTLTCHEVVELTEDGPRPOQSSNMHNSHLPYLLSSQMSKRTYQSGFHNKNT	415
Db	361	ornlftltechvveltedlgprpqssnmhns1pctalsesqmsrtyqsgfhnknt	415

RESULT	5
AAE03629	
ID	AAE03629 standard; Protein; 415 AA.
XX	
AC	AAE03629;
XX	
DT	07-AUG-2001 (first entry)
XX	
DE	Human G-protein coupled receptor, SNORF72.
XX	
KW	Human; G-protein coupled receptor; SNORF72; neuropeptide;
KW	NMD; inflammation; arthritis; autoimmune disease; septicemia; psychotic
KM	mental retardation; transplant rejection; schizophrenia; dementia; obesity; pain
KM	respiratory disorder; depression; hypotension; epilepsy; diabetes
KM	gastrointestinal disorder; hypertension; circadian disorder; anorexia
KM	schmeia; stroke; cancer; sexual disorder; bulimia; allergy
KM	neurological; psoriasis; Parkinson's disease; nausea; asthma;
KM	Almer's disease; AIDS; hormonal disorder; memory disorder; migraine;
KM	vascular disorder; renal disorder; bone disease; delirium; asthma;

XX	Homo sapiens.
OS	
XX	
XX	Key
FH	Location/Qualifiers
FT	Domain
FT	/label = Transmembrane_domain
FT	82..98
FT	/label = Transmembrane_domain
FT	124..142
FT	/label = Transmembrane_domain
FT	166..179
FT	/label = Transmembrane_domain
FT	215..240
FT	/label = Transmembrane_domain
FT	269..286
FT	/label = Transmembrane_domain
FT	301..330
FT	/label = Transmembrane_domain
PN	WO200144297-A1.
XX	
XX	21-JUN-2001.
XX	
PF	13-DEC-2000; 2000MO-US33787.
PR	17-DEC-1999; 99US-0466435.
PR	25-APR-2000; 2000US-0558099.
PR	30-JUN-2000; 2000US-0609146.
PA	(SYNA-) SYNAPTIC PHARM CORP.
PI	Bonini JA, Lerman GS, Quan Y, Gogozalek K;
DR	WPI: 2001-390240/A1.
N-PSDB:	AAD08008.
P7	A purified mammalian SNORF62 or SNORF72 receptor protein for
PT	identification of compounds to treat e.g. inflammation, arthritis,
PT	autoimmune diseases, transplant rejection, AIDS, cancer -
XX	
PS	Claim 12; Fig 4; 256pp; English.
XX	
CC	The invention relates to human G-protein coupled receptors, SNORF62 and
CC	SNORF72 and their corresponding cDN molecules. SNORF62 and SNORF72
CC	receptors are specific for neuromedin U (NMU) neuropeptides, hence they
CC	are also known as NMU receptors. The agonist and antagonist of NMU
CC	receptors are useful for treating an abnormality in a subject that is
CC	alleivated by decreasing or increasing the activity of NMU receptor.
CC	The NMU receptors serves as a valuable tool for designing drugs which are
CC	useful for treating various pathophysiological conditions such as
CC	Inflammation, arthritis, autoimmune diseases, transplant rejection,
CC	graft vs host disease, bacterial, fungal, protozoan and viral infections,
CC	septicemia, AIDS, pain, psychotic and neurological disorders, including
CC	anxiety, depression, schizophrenia, dementia, mental retardation, memory
CC	loss, epilepsy, neuromotor disorders, respiratory disorders, asthma,
CC	eating/body weight disorders including obesity, bulimia, diabetes,
CC	anorexia, nausea, hypertension, hypotension, vascular and cardiovascular
CC	disorders, renal disorders, bone diseases including osteoporosis, benign
CC	prostatic hypertrophy, gastrointestinal disorders, nasal congestion,
CC	dermatological disorders such as psoriasis, allergies, Parkinson's
CC	disease, Alzheimer's disease, acute heart failure, angina disorders, etc
CC	delirium and dyskinesias such as Huntington's disease. They can also be
CC	used to regulate steroid hormone disorders, endocrine disorders, memory
CC	disorders, electrolyte balance disorders, endocrine disorders, behavioral
CC	disorders, somatosensory disorders, migraine, Addison's disease, Cushing's
CC	disease, prevent miscarriage, induce labour or to treat dysmenorrhea.
CC	The present sequence is human G-protein coupled receptor, SNORF72.
XX	

SQ Sequence 415 AA:

Query Match 100.0%: Score 2185; DB 22; Length 415;
Best Local Similarity 100.0%: Pred. No. 9.4e-246;
Matches 415; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MSGMEKLAQNASWYQOKLEDPFOKHLNSTREXYLAFLCGPRSHFFLPVSVVVPPIFVGV 60
DB 1 msgmeklqnaswlyqkledpfqkhnstceeylaficgprshfflpvsvvvpplfvgv 60
OY 61 IGVNLVCLVILQHOAKMPTNTYLLSLAVSDLLVLLGMPLEVEWMRYPELFGVGCY 120
DB 61 IGVNLVCLVILQHOAKMPTNTYLLSLAVSDLLVLLGMPLEVEWMRYPELFGVGCY 120
OY 121 FKTALEFETWCFSALSTITVSVERVYALHPRAKIQSTRRALRLGLVWGFVLSLP 180
DB 121 FKTALEFETWCFSALSTITVSVERVYALHPRAKIQSTRRALRLGLVWGFVLSLP 180
OY 181 NTSIHGIRKHPNGSLVPGSATCTVYIKPMWYNTFIQTSFLEYLLPMTVISLYLMA 240
DB 181 ntsihgirkhpnngslvpgsatctv yikpmw yntfiqtsfleyllpmtvislylma 240
OY 241 LRLKDKSLAEADGNANIORPCRSYKMLFVLVLFALCMAFPHIDRLFSFVEWSES 300
DB 241 LRLKDKSLAEADGNANIORPCRSYKMLFVLVLFALCMAFPHIDRLFSFVEWSES 300
OY 301 LAAVFNLVHVSGVFYSSAVNPITYNLSRFOAFQNVNSSFHKNHSHQHDPLPA 360
DB 301 laavfnlvhvsgvf yssavnpit ynlssrfoafqnv nssf hknhs hqhdplpa 360
OY 361 QRIIFLTECHFVELMEDIGPOPCOSSMHNHSLPTALSSEQMSRTVYQSFHNKT 415
DB 361 qriifltechfvelmedigpofpcossmhnhslptalsseqmsrtv yqsfh nkt 415

RESULT 6
AAB67802
ID AAB67802 standard; Protein: 415 AA.

AC AAB67802;
DT 29-JUN-2001 (first entry)
DE Amino acid sequence of G-protein coupled receptor IGSA4 long version.
KW Human; G-protein coupled receptor; IGSA4; IGSA4A; IGSA4B; schizophrenia;
KW nervous system disorder; psychiatric disorder; Parkinson's disease;
KW episodic paroxysmal anxiety; phobia; migraine; bulimia; stroke;
KW cardiovascular disease; heart failure; angina pectoris; obesity; emesis;
KW motility disorder; myocardial infarction; hypertension; dyslipidemia;
KW gastrointestinal disorder; inflammatory bowel disease; osteoporosis;
KW inflammation; infection; pain; cancer; immune disorder; allergy; sepsis;
KW gynecological disorder.
OS Homo sapiens.
PN WO200125269-A2.
PD 12-APR-2001.
PF 25-SEP-2000; 2000WO-EP09584.
PR 24-SEP-1999; 99EP-0203140.
PR 24-SEP-1999; 99NL-1013140.
PR 28-JUL-2000; 2000EP-0202683.
PR 31-JUL-2000; 2000US-0222047.
PA (SOLV) SOLVAY PHARM BV.
PI Deleerenlijder W, Berger C, Loeken C, Nys G, Venema J;
XX WPI: 2001-273568/28.

DR N-PSDB: AAF80322.

PT New G-protein coupled receptors and the polynucleotides encoding them,
PT useful for preventing, ameliorating or correcting nervous system
PT disorders, cardiovascular diseases, dyslipidemias, inflammations, pain
or cancers -

PS Claim 18; Page 77-79; 102pp; English.

CC The present sequence represents the long version of a human G-protein
CC coupled receptor designated IGSA4. IGSA4 exists in two polymorphic forms,
CC IGSA4 and IGSA4B. The IGSA4 receptors and IGSA4 polynucleotides are useful
CC for preventing, ameliorating or correcting dysfunction or diseases.
CC These diseases include peripheral nervous system, psychiatric and central
CC nervous system disorders (e.g. schizophrenia, episodic paroxysmal
CC anxiety, phobia, Parkinson's disease, migraine, epilepsy, bulimia or
CC stroke), cardiovascular diseases (e.g. heart failure, angina pectoris,
CC myocardial infarction or hypertension), dyslipidemias, obesity, emesis,
CC gastrointestinal disorders (e.g. inflammatory bowel disease or motility
CC disorders), osteoporosis, inflammations, infections (e.g. bacterial,
CC fungal, protozoan or viral), pain, cancers, immune disorders, allergies,
CC sepsis or gynecological disorders. Agonists or antagonists of IGSA4 are
CC effective with regard to disorders of the nervous system, including the
CC central and peripheral nervous systems, disorders of the gastrointestinal
CC system, cardiovascular system, skeletal muscle, thyroid, lung or
CC genitourinary system, or immunological disease. The IGSA4 polynucleotides
CC are useful as diagnostic reagents for detecting under-expression,
CC overexpression or altered expression of IGSA4.

SQ Sequence 415 AA:

Query Match 100.0%: Score 2185; DB 22; Length 415;
Best Local Similarity 100.0%: Pred. No. 9.4e-246;
Matches 415; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MSGMEKLAQNASWYQOKLEDPFOKHLNSTREXYLAFLCGPRSHFFLPVSVVVPPIFVGV 60
DB 1 msgmeklqnaswlyqkledpfqkhnstceeylaficgprshfflpvsvvvpplfvgv 60
OY 61 IGVNLVCLVILQHOAKMPTNTYLLSLAVSDLLVLLGMPLEVEWMRYPELFGVGCY 120
DB 61 IGVNLVCLVILQHOAKMPTNTYLLSLAVSDLLVLLGMPLEVEWMRYPELFGVGCY 120
OY 121 FKTALEFETWCFSALSTITVSVERVYALHPRAKIQSTRRALRLGLVWGFVLSLP 180
DB 121 FKTALEFETWCFSALSTITVSVERVYALHPRAKIQSTRRALRLGLVWGFVLSLP 180
OY 181 NTSIHGIRKHPNGSLVPGSATCTVYIKPMWYNTFIQTSFLEYLLPMTVISLYLMA 240
DB 181 ntsihgirkhpnngslvpgsatctv yikpmw yntfiqtsfleyllpmtvislylma 240
OY 241 LRLKDKSLAEADGNANIORPCRSYKMLFVLVLFALCMAFPHIDRLFSFVEWSES 300
DB 241 LRLKDKSLAEADGNANIORPCRSYKMLFVLVLFALCMAFPHIDRLFSFVEWSES 300
OY 301 LAAVFNLVHVSGVFYSSAVNPITYNLSRFOAFQNVNSSFHKNHSHQHDPLPA 360
DB 301 laavfnlvhvsgvf yssavnpit ynlssrfoafqnv nssf hknhs hqhdplpa 360
OY 361 QRIIFLTECHFVELMEDIGPOPCOSSMHNHSLPTALSSEQMSRTVYQSFHNKT 415
DB 361 qriifltechfvelmedigpofpcossmhnhslptalsseqmsrtv yqsfh nkt 415

RESULT 7
AAB67803
ID AAB67803 standard; Protein: 412 AA.

AC AAB67803;
DT 29-JUN-2001 (first entry)
XX

DE Amino acid sequence of G-protein coupled receptor IGS4A short version.
 XX
 XX Human: G-protein coupled receptor: IGS4: IGS4A: IGS4B: schizophrenia:
 KM nervous system disorder; psychiatric disorder; Parkinson's disease;
 KM episodic paroxysmal anxiety; phobia; migraine; epilepsy; bulimia; stroke;
 KM cardiovascular disease; heart failure; angina pectoris; obesity; emesis;
 KM motility disorder; myocardial infarction; hypertension; dyslipidemia;
 KM gastrointestinal disorder; inflammatory bowel disease; osteoporosis;
 KM inflammation; infection; pain; cancer; immune disorder; allergy; sepsis;
 KM gynecological disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200125269-A2.
 XX
 PD 12-APR-2001.
 XX
 PF 25-SEP-2000; 2000MO-EP09584.
 XX
 PR 24-SEP-1999; 99EP-0203140.
 XX 24-SEP-1999; 99NL-1013140.
 PR 28-JUL-2000; 2000EP-0202683.
 PR 31-JUL-2000; 2000US-0222047.
 XX
 PA (SOLV) SOLVAY PHARM BV.
 XX
 PI Deleersnyder W, Berger C, Loeken C, Nys G, Venema J;
 DR WPI, 2001-273568/28.
 DR N-PSDB; AAFF0323.
 XX
 PT New G-protein coupled receptors and the polynucleotides encoding them,
 PT useful for preventing, ameliorating or correcting nervous system
 PT disorders, cardiovascular diseases, dyslipidemias, inflammations, pain
 PT or cancers
 XX
 PS Claim 18: Page 81-82; 102pp; English.
 XX
 CC The present sequence represents the short version of a human G-protein
 CC coupled receptor designated IGS4A. IGS4 exists in two polymorphic forms,
 CC IGS4A and IGS4B. The IGS4 receptors and IGS4 polynucleotides are useful
 CC for preventing, ameliorating or correcting dysfunctions or diseases.
 CC These diseases include peripheral nervous system, psychiatric and central
 CC nervous system disorders (e.g. schizophrenia, episodic paroxysmal
 CC anxiety, phobia, Parkinson's disease, migraine, epilepsy, bulimia or
 CC stroke), cardiovascular diseases (e.g. heart failure, angina pectoris,
 CC myocardial infarction or hypertension), dyslipidemias, obesity, emesis,
 CC gastrointestinal disorders (e.g. inflammatory bowel disease or motility
 CC disorders), osteoporosis, inflammations, infections (e.g. bacterial,
 CC fungal, protozoan or viral), pain, cancers, immune disorders, allergies,
 CC sepsis or gynecological disorders. Agonists or antagonists of IGS4 are
 CC effective with regard to disorders of the nervous system, including the
 CC central and peripheral nervous systems, disorders of the gastrointestinal
 CC system, cardiovascular system, skeletal muscle, thyroid, lung or
 CC genitourinary system, or immunological disease. The IGS4 polynucleotides
 CC are useful as diagnostic reagents for detecting under-expression,
 CC overexpression or altered expression of IGS4.
 CC
 XX
 SQ Sequence 412 AA:

Query Match 99.3%; Score 2170; DB 22; Length 412;
 Best Local Similarity 100.0%; Pred. No. 5.2e-244;
 Matches 412; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 MEKLNQNSWITQOKLEDFQKHLNSTEYLAFLCGPRRSHFLPVSVVYPIFVVGIGN 63
 DB 1 MEKLNQNSWITQOKLEDFQKHLNSTEYLAFLCGPRRSHFLPVSVVYPIFVVGIGN 60
 OY 64 VIVCLVTLQHQAMKTPNNYVLFSLAVSDLVLLGLMPLFVEMKRNRPFLGPGVCYEKT 123
 DB 61 VIVCLVTLQHQAMKTPNNYVLFSLAVSDLVLLGLMPLFVEMKRNRPFLGPGVCYEKT 120

OY 124 ALFEIVCFASLISITTVSVERYVALHPERAKLQSTRRARILIGVGFSLFSPNTS 183
 DB 121 ALFEIVCFASLISITTVSVERYVALHPERAKLQSTRRARILIGVGFSLFSPNTS 180
 OY 184 IRIKIRHYPNGSLVGSATCTYIKRMWITNFIQVTSFLFYLIPMTVLSVLYLMAAL 243
 DB 181 IRIKIRHYPNGSLVGSATCTYIKRMWITNFIQVTSFLFYLIPMTVLSVLYLMAAL 240
 OY 244 KKDLSLEADGNNIOPKCRKSVKMLFVLVFAICMPPHIDRLFFSFVEESESLLA 303
 DB 241 KKDLSLEADGNNIOPKCRKSVKMLFVLVFAICMPPHIDRLFFSFVEESESLLA 300
 OY 304 VFNLVHVSGVFYSSAVNPIIYNLISRRFQAQFQNTVSSFHQWNSQHDPLPQAQN 363
 DB 301 VFNLVHVSGVFYSSAVNPIIYNLISRRFQAQFQNTVSSFHQWNSQHDPLPQAQN 360
 OY 364 IFLTECHFVBLTDIGPQPCQSSMHNHSLPTALSSQKSRNTYQSFHNKT 415
 DB 361 IFLTECHFVBLTDIGPQPCQSSMHNHSLPTALSSQKSRNTYQSFHNKT 412

RESULT 8

AA63366 standard; Protein; 415 AA.
 XX
 AC AA63366;
 XX
 DT 15-OCT-2001 (first entry)
 XX
 DE Amino acid sequence of a human TGR-1 protein.
 XX
 KM TGR-1; neuromedin U; hypertension; stress disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200157524-A1.
 XX
 PD 09-AUG-2001.
 XX
 PF 02-FEB-2001; 2001MO-JP00746.
 XX
 PR 04-FEB-2000; 2000JP-0032773.
 PR 24-FEB-2000; 2000JP-0052252.
 PR 30-MAR-2000; 2000JP-0097896.
 PR 19-JUN-2000; 2000JP-0187536.
 XX
 PA (TAKE) TAKEDA CHEM IND LTD.
 XX
 PI Hinuma S, Shintani Y, Hosoya M, Fujii R, Moriya T, Matsui H;
 PI Okubo S;
 DR WPI: 2001-488917/53.
 DR N-PSDB; AAH43075.
 XX
 PT Identifying predicted or actual structures of two or more members of a
 PT chemical or physical library by mass spectrometry comprising
 PT correlating molecular mass measurements of two or more members with a
 PT shared chemical history -
 XX
 PS disclosure; Page 88; 95pp; Japanese.
 XX
 CC The present sequence represents a human TGR-1 protein. The specification
 CC describes a method of screening a compound, which is capable of binding
 CC properties of neuromedin U to TGR-1. The method is useful for screening
 CC preventatives and remedies for hypertension, stress diseases, etc..
 CC TGR-1 antagonists are also useful for treating the same diseases.
 CC
 XX
 SQ Sequence 415 AA:

Query Match 99.1%; Score 2166; DB 22; Length 415;
 Best Local Similarity 99.0%; Pred. No. 1.5e-243;
 Matches 411; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSGMERLQNASWYIOOKLEDPEOKHLNSTEELYLAFLCGPRSHFFLPVSVVYPIFVGV 60
DB 1 msgmeklqnaswlyyqgkledpfqkhnstceeylaficgprshfflpvsvvypifvgy 60
QY 61 IGVNLVCLVILQHOAMKPTNTYLFSLAVSDLLVLLGMPLEVEYEMWRNYPFLFGVGCY 120
DB 61 IGVNLVCLVILQHOAMKPTNTYLFSLAVSDLLVLLGMPLEVEYEMWRNYPFLFGVGCY 120
QY 121 FRTALFETVCFASILSTTVSVERYVAIILHPPRAKIQSTRRRALRLIGYWGSSVLFSLP 180
DB 121 FRTALFETVCFASILSTTVSVERYVAIILHPPRAKIQSTRRRALRLIGYWGSSVLFSLP 180
QY 181 NTSIHGKIFKFPNGSLVPGSATCTVIRKPMWYINFIQVTSFLFYLLPMTVIVLYLMA 240
DB 181 ntsihgkifkfyfngslvpgsatctvirkpmwylfnfiqvtsflfyllpmtvisvlylma 240
QY 241 LRLKKDLSLEADGNNANIQPCRKSVNKMFLVILVFAICWAPFHIDRLFSSVEWSES 300
DB 241 LRLKKDLSLEADGNNANIQPCRKSVNKMFLVILVFAICWAPFHIDRLFSSVEWSES 300
QY 301 LAAVFNLVHVSGVFYLSAVNPITYNLSRFOAFOVNISSFKOMHSQHDPLPPA 360
DB 301 laavfnlvhvsgvfyllsavnpitynlsrfoafovniissfkomhsqhdplppa 360
QY 361 ORNIFLTECHFVELTEDIGPQPCOSSMHNHSLPTALSSQMSRTNYQSFFHFKT 415
DB 361 grnifltechfveltedigpqpccssmhnshlptalsseqmstrnyqsfhfknt 415

RESULT 9
AAB67804
ID AAB67804 standard; Protein; 415 AA.

AC AAB67804;

DT 29-JUN-2001 (first entry)

XX Amino acid sequence of a G-protein coupled receptor IG54B long version.

Human; G-protein coupled receptor; IG54; IG54A; IG54B; schizophrenia;
nervous system disorder; psychiatric disorder; Parkinson's disease;
episodic paroxysmal anxiety; phobia; migraine; epilepsy; bulimia; stroke;
cardiovascular disease; heart failure; angina pectoris; obesity; emesis;
motility disorder; myocardial infarction; hypertension; dyslipidemia;
gastrointestinal disorder; inflammatory bowel disease; osteoporosis;
inflammation; infection; pain; cancer; immune disorder; allergy; sepsis;
gynecological disorder.

OS Homo sapiens.

PN WO200125269-A2.

PD 12-APR-2001.

PE 25-SEP-2000; 2000WO-EP09584.

PR 24-SEP-1999; 99EP-0203140.

PR 24-SEP-1999; 99WL-1013140.

PR 28-JUL-2000; 2000EP-0202683.

PR 31-JUL-2000; 2000US-0222047.

PA (SOLV) SOLVAY PHARM BV.

PI Deleersnijder W, Berger C, Loeken C, Nys G, Venema J;

DR N-PSDB; AAF80324.

XX New G-protein coupled receptors and the polynucleotides encoding them,
PT useful for preventing, ameliorating or correcting nervous system
PT disorders, cardiovascular diseases, dyslipidemias, inflammations, pain
PT or cancers

XX Claim 19; Page 85-86; 102pp; English.

CC The present sequence represents the long version of a human G-protein
CC coupled receptor designated IG54B. IG54 exists in two polymorphic forms,
CC IG54A and IG54B. The IG54 receptors and IG54 polynucleotides are useful
CC for preventing, ameliorating or correcting dysfunctions or diseases.
CC These diseases include peripheral nervous system, psychiatric and central
CC nervous system disorders (e.g. schizophrenia, episodic paroxysmal
CC anxiety, phobia, Parkinson's disease, migraine, epilepsy, bulimia or
CC stroke), cardiovascular diseases (e.g. heart failure, angina pectoris,
CC myocardial infarction or hypertension), dyslipidemias, obesity, emesis,
CC gastrointestinal disorders (e.g. inflammatory bowel disease or motility
CC disorders), osteoporosis, inflammations, infections (e.g. bacterial,
CC fungal, protozoan or viral), pain, cancers, immune disorders, allergies,
CC sepsis or gynecological disorders. Agonists or antagonists of IG54 are
CC effective with regard to disorders of the nervous system, including the
CC central and peripheral nervous systems, disorders of the gastrointestinal
CC system, cardiovascular system, skeletal muscle, thyroid, lung or
CC genitourinary system, or immunological disease. The IG54 polynucleotides
CC are useful as diagnostic reagents for detecting under-expression,
CC overexpression or altered expression of IG54.

SQ Sequence 415 AA;

Query Match 98.9%; Score 2162; DB 22; Length 415;
Best Local Similarity 99.08; Pred. No. 4,5e-243;
Matches 411; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSGMERLQNASWYIOOKLEDPEOKHLNSTEELYLAFLCGPRSHFFLPVSVVYPIFVGV 60
DB 1 msgmeklqnaswlyyqgkledpfqkhnstceeylaficgprshfflpvsvvypifvgy 60
QY 61 IGVNLVCLVILQHOAMKPTNTYLFSLAVSDLLVLLGMPLEVEYEMWRNYPFLFGVGCY 120
DB 61 IGVNLVCLVILQHOAMKPTNTYLFSLAVSDLLVLLGMPLEVEYEMWRNYPFLFGVGCY 120
QY 121 FRTALFETVCFASILSTTVSVERYVAIILHPPRAKIQSTRRRALRLIGYWGSSVLFSLP 180
DB 121 frtalfetvcafslstttvsveryvaailhpprakiqstrrralrligywgsylvfslp 180
QY 181 NTSIHGKIFKFPNGSLVPGSATCTVIRKPMWYINFIQVTSFLFYLLPMTVIVLYLMA 240
DB 181 ntsihgkifkfyfngslvpgsatctvirkpmwylfnfiqvtsflfyllpmtvisvlylma 240
QY 241 LRLKKDLSLEADGNNANIQPCRKSVNKMFLVILVFAICWAPFHIDRLFSSVEWSES 300
DB 241 LRLKKDLSLEADGNNANIQPCRKSVNKMFLVILVFAICWAPFHIDRLFSSVEWSES 300
QY 301 LAAVFNLVHVSGVFYLSAVNPITYNLSRFOAFOVNISSFKOMHSQHDPLPPA 360
DB 301 laavfnlvhvsgvfyllsavnpitynlsrfoafovniissfkomhsqhdplppa 360
QY 361 ORNIFLTECHFVELTEDIGPQPCOSSMHNHSLPTALSSQMSRTNYQSFFHFKT 415
DB 361 grnifltechfveltedigpqpccssmhnshlptalsseqmstrnyqsfhfknt 415

RESULT 10
AAV52992
ID AAV52992 standard; Protein; 415 AA.

AC AAV52992;

DT 21-FEB-2000 (first entry)

XX Human neurotensin-like receptor.

KW Human; neurotensin-like receptor; NLR: G-protein coupled receptor;

OS central nervous system; analgesia.

XX Homo sapiens.

XX MO9955732-A1.
 XX 04-NOV-1999.
 XX 15-APR-1999; 99WO-SE00598.
 XX 24-APR-1998; 98SE-0001455.
 XX (ASTR-) ASTRA PHARMA INC.
 XX (ASTR) ASTRA AB.
 PI Ahmed S, Cao J, O'Donnell D, Walker P;
 XX WPI: 2000-052803/04.
 DR N-PSDB; AA233297.
 XX Novel neurotensin-like receptor, useful for identifying agents for
 PT producing anesthesia or analgesia -
 PS Claim 1; Fig 2; 46pp; English.
 XX The present sequence represents a human G-protein coupled receptor
 CC neurotensin-like receptor (NLR). The NLR polynucleotide and protein
 CC can be used to isolate compounds that bind, (ant)agonise or alter
 CC the activity or expression of the NLR. The NLR is a G-protein coupled
 CC receptor which is expressed in the central nervous system and shares
 CC homology with human neurotensin receptor. The receptors can be used in
 CC assays to identify agents for producing anesthesia and analgesia.
 XX Sequence 415 AA:

Query Match 98.3%; Score 2148; DB 21; Length 415;
 Best Local Similarity 98.6%; Pred. No. 1.9e-241;
 Matches 409; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 QY 1 MSGMEKLNASWITQOKLEDPFQKHLNSTEYLAFLCGPRSRHFFLVSVVYVPFVGV 60
 DB 1 msgmeklnaswitygkledpfqkhlNSTEYLAFLCGPRSRHFFLVSVVYVPFVGV 60
 QY 61 IGNVLVCVLIQHOAMKPTNYVYFSLAVSDLVLLGMPLEVEYEMRNVPFLGPGVCY 120
 DB 61 IGNVLVCVLIQHOAMKPTNYVYFSLAVSDLVLLGMPLEVEYEMRNVPFLGPGVCY 120
 QY 121 FKTLFETVCASLITSTTVSEYRYVALHPFRAKLOSTRRRARILGIVGFSYLSLP 180
 DB 121 FKTLFETVCASLITSTTVSEYRYVALHPFRAKLOSTRRRARILGIVGFSYLSLP 180
 QY 121 FKTLFETVCASLITSTTVSEYRYVALHPFRAKLOSTRRRARILGIVGFSYLSLP 180
 DB 121 FKTLFETVCASLITSTTVSEYRYVALHPFRAKLOSTRRRARILGIVGFSYLSLP 180
 QY 181 NTSIHGIRKHFPGSLVPGSATCTVTKPMYINFTIOVTSFLYLLPMYISVLYYMA 240
 DB 181 ntsihgirkhfpgslvpgsatctvtkpmYINFTIOVTSFLYLLPMYISVLYYMA 240
 QY 241 LRLKDSLEDEGNANIORCGRSVNKMFLVLLVFAICAPPHIDLFSPFEWSES 300
 DB 241 LRLKDSLEDEGNANIORCGRSVNKMFLVLLVFAICAPPHIDLFSPFEWSES 300
 QY 301 LAAVFNLVHVVSGVFYLLSAVNPIYVNLRSRFOAFQVNISSPHKQMSQHPOLPPA 360
 DB 301 LAAVFNLVHVVSGVFYLLSAVNPIYVNLRSRFOAFQVNISSPHKQMSQHPOLPPA 360
 QY 361 ORNIFLECHVVELTEDIGPOPCOSSMHNHSLPVALSSEDMSTRNQSHPNKT 415
 DB 361 qrniflechvveltedigpocossMHNHSLPVALSSEDMSTRNQSHPNKT 415

RESULT 11
 AAB67805
 ID AAB67805 standard; Protein: 412 AA.
 XX AAB67805;
 AC
 XX 29-JUN-2001 (first entry)

XX Amino acid sequence of G-protein coupled receptor IGS4B short version.
 DE Human: G-protein coupled receptor; IGS4; IGS4B; schizophrenia;
 XX nervous system disorder; psychiatric disorder; Parkinson's disease;
 KW episodic paroxysmal anxiety; phobia; migraine; epilepsy; bulimia; stroke;
 KW cardiovascular disease; heart failure; angina pectoris; obesity; emesis;
 KW mobility disorder; myocardial infarction; hypertension; dyslipidemia;
 KW gastrointestinal disorder; inflammatory bowel disease; osteoporosis;
 KW inflammation; infection; immune disorder; cancer; immune disorder; allergy; sepsis;
 KW gynecological disorder.
 XX Homo sapiens.
 OS
 XX WO200125269-A2.
 XX 12-APR-2001.
 XX 25-SEP-2000; 2000WO-EP09584.
 XX 24-SEP-1999; 99EP-0203140.
 PR 24-SEP-1999; 99NL-1013140.
 PR 28-JUL-2000; 2000EP-0202683.
 PR 31-JUL-2000; 2000US-0222047.
 XX (SOLV) SOLVAY PHARM BV.
 PA
 PI Deleersnyder W, Berger C, Loeken C, Nys G, Venema J;
 XX WPI: 2001-273568/28.
 DR N-PSDB; AAF80325.
 XX New G-protein coupled receptors and the polynucleotides encoding them,
 PT useful for preventing, ameliorating or correcting nervous system
 PT disorders, cardiovascular diseases, dyslipidemias, inflammations, pain
 PT or cancers -
 PS Claim 18; Page 89-90; 102pp; English.
 XX The present sequence represents the short version of a human G-protein
 CC coupled receptor designated IGS4B. IGS4 exists in two polymorphic forms,
 CC IGS4A and IGS4B. The IGS4 receptors and IGS4 polynucleotides are useful
 CC for preventing, ameliorating or correcting dysfunctions or diseases.
 CC These diseases include peripheral nervous system, psychiatric and central
 CC nervous system disorders (e.g. schizophrenia, episodic paroxysmal
 CC anxiety, phobia, Parkinson's disease, migraine, epilepsy, bulimia or
 CC stroke), cardiovascular diseases (e.g. heart failure, angina pectoris,
 CC myocardial infarction or hypertension), dyslipidemias, obesity, emesis,
 CC gastrointestinal disorders (e.g. inflammatory bowel disease or mobility
 CC disorders), osteoporosis, inflammations, infections (e.g. bacterial,
 CC fungal, protozoan or viral), pain, cancers, immune disorders, allergies,
 CC sepsis or gynecological disorders. Agonists or antagonists of IGS4 are
 CC effective with regard to disorders of the nervous system, including the
 CC central and peripheral nervous systems, disorders of the gastrointestinal
 CC system, cardiovascular system, skeletal muscle, thyroid, lung or
 CC genitourinary system, or immunological disease. The IGS4 polynucleotides
 CC are useful as diagnostic reagents for detecting under-expression,
 CC overexpression or altered expression of IGS4.
 XX Sequence 412 AA:

Query Match 98.3%; Score 2147; DB 22; Length 412;
 Best Local Similarity 99.0%; Pred. No. 2.5e-241;
 Matches 408; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 4 MEKLNASWITQOKLEDPFQKHLNSTEYLAFLCGPRSRHFFLVSVVYVPFVGVGN 63
 DB 1 meklnaswitygkledpfqkhlNSTEYLAFLCGPRSRHFFLVSVVYVPFVGVGN 60
 QY 64 VLVCLVLIQHOAMKPTNYVYFSLAVSDLVLLGMPLEVEYEMRNVPFLGPGVCYFKT 123
 DB 61 vlvclvliqhoamkptnyyflsLAVSDLVLLGMPLEVEYEMRNVPFLGPGVCYFKT 120

OY 124 ALFETVCFASILSTTVERVAAILHPFRALQSTRRALRILGIWGSVLSLEPNTS 183
 DB 121 alfetvcfasilsttvervvaailhpfraklqstrralrllglwgsvlsfslpnts 180
 OY 184 IHGKHFYFPNGSLVPSGATCTVYIKPMWYITFIOTSEFLYLLPMVITVLYLMAURL 243
 DB 181 ihgkfhfypngslvpsgactctvlykpmwlyitfioyseflyllpmtvlslylmrl 240
 OY 244 KKKSLEADGGMANIOPRCRSYVNMFLVLYFAICWAPFHIDRLEFSVEESLAA 303
 DB 241 kksleadeegmanioqrctsvnmflvlyfalcwaphidrlfseveeslaa 300
 OY 304 VFNLVHVSGVFYFLSSAVNPITYNLRSRFOAFQNVISFHKOMHSQHPOLPPAQRN 363
 DB 301 vfnlvhvsgvfyflssavnpilynlrsrfqafqnvlsfkhqmsqhdppqpqrn 360
 OY 364 IFLTECHFEVLTDIGPOPCQSSMHNHLPTALSSQMSRTNYQSFHNKT 415
 DB 361 ifltechfevltdigpqcqssmhnshlptalsseqmstrnyqsfhnkt 412
 RESULT 12
 ID AAE03634 standard; Protein: 395 AA.
 XX AAE03634;
 AC AAE03634;
 XX 07-AUG-2001 (first entry)
 DE Rat G-protein coupled receptor, SNORF72.
 XX
 KM Rat: G-protein coupled receptor; SNORF72; neuromedin U neuropeptide;
 KM MMU; inflammation; arthritis; autoimmune disease; septicemia; psychotic;
 KM mental retardation; transplant rejection; neurological disorder; anxiety;
 KM respiratory disorder; depression; schizophrenia; dementia; obesity; pain;
 KM gastrointestinal disorder; hypertension; hypotension; epilepsy; diabetes;
 KM ischemia; stroke; cancer; sexual disorder; circadian disorder; anorexia;
 KM dermatological; psoriasis; Parkinson's disease; nausea; bulimia; allergy;
 KM Alzheimer's disease; AIDS; hormonal disorder; memory disorder; migraine;
 KM cardiovascular disorder; renal disorder; bone disease; delirium; asthma;
 KM Cushing's disease; dysmenorrhea; antidiabetic; osteoporosis;
 KM metabolic disorder; behavioral disorder; Addison's disease; dyskinesia;
 KM tranquilizer; antiulcer; antiaddictive.
 XX
 OS Rattus sp.
 FH
 FH Key location/Qualifiers
 FT Domain 40..63
 FT Domain /label= Transmembrane_domain
 FT Domain 75..95
 FT Domain /label= Transmembrane_domain
 FT Domain 106..132
 FT Domain /label= Transmembrane_domain
 FT Domain 156..175
 FT Domain /label= Transmembrane_domain
 FT Domain 202..227
 FT Domain /label= Transmembrane_domain
 FT Domain 259..278
 FT Domain /label= Transmembrane_domain
 FT Domain 292..322
 FT Domain /label= Transmembrane_domain
 PN MO200144297-A1.
 XX 21-JUN-2001.
 PF 13-DEC-2000; 2000WO-US33787.
 XX 17-DEC-1999; 99US-0466435.
 PR 25-APR-2000; 2000US-0558099.
 PR 30-JUN-2000; 2000US-0609146.

PA (SYNA-) SYNAPTIC PHARM CORP.
 XX Bonini JA, Lerman GS, Quan Y, Ogozalek K;
 PI WPI, 2001-390240/41.
 DR N-PSDB: AAD08024.
 XX
 PT A purified mammalian SNORF62 or SNORF72 receptor protein for
 PT identification of compounds to treat e.g. inflammation, arthritis,
 PT autoimmune diseases, transplant rejection, AIDS, cancer -
 PS Claim 15; Fig 15; 256pp; English.
 XX
 CC The invention relates to human G-protein coupled receptors, SNORF62 and
 CC SNORF72 and their corresponding cDNA molecules. SNORF62 and SNORF72
 CC receptors are specific for neuromedin U (NMU) neuropeptides, hence they
 CC are also known as NMU receptors. The agonist and antagonist of NMU
 CC receptors are useful for treating an abnormality in a subject that is
 CC alleviated by decreasing or increasing the activity of NMU receptor.
 CC The NMU receptors serves as a valuable tool for designing drugs which are
 CC useful for treating various pathological conditions such as
 CC inflammation, arthritis, autoimmune diseases, transplant rejection,
 CC graft vs host disease, bacterial, fungal, protozoan and viral infections,
 CC septicemia, AIDS, pain, psychotic and neurological disorders, including
 CC anxiety, depression, schizophrenia, dementia, mental retardation, memory
 CC loss, epilepsy, neuromotor disorders, respiratory disorders, asthma,
 CC eating/body weight disorders including obesity, bulimia, diabetes,
 CC anorexia, nausea, hypertension, hypotension, vascular and cardiovascular
 CC disorders, ischemia, stroke, cancers, sexual disorders, circadian
 CC disorders, renal disorders, bone diseases including osteoporosis, benign
 CC prostatic hypertrophy, gastrointestinal disorders, nasal congestion,
 CC dermatological disorders such as psoriasis, allergies, Parkinson's
 CC disease, Alzheimer's disease, acute heart failure, angina disorders,
 CC delirium and dyskinesias such as Huntington's disease. They can also be
 CC used to regulate steroid hormone disorders, epinephrine release
 CC disorders, electrolyte balance disorders, endocrine disorders, memory
 CC disorders, somatosensory balance disorders, metabolic disorders, behavioral
 CC disorders, drug addiction, migraine, Addison's disease, Cushing's
 CC disease, prevent miscarriage, induce labour or to treat dysmenorrhea.
 CC The present sequence is rat G-protein coupled receptor, SNORF72 related
 CC to the invention.
 CC
 XX
 XX Sequence 395 AA;
 SQ
 Query Match 75.58; Score 1650.5; DB 22; Length 395;
 Best Local Similarity 79.44; Pred. No. 176-183;
 Matches 312; Conservative 36; Mismatches 40; Indels 5; Gaps 1;
 OY 4 MEKQNASWITYOQLEPPFOKHLNSTEEYLAFLGPRSHFFLPSVVPYIPVGYIGN 63
 DB 1 mglknsawih-----dplmkylnsteeylahlcgprrsdlsprsvayallilivgyign 55
 OY 64 VLQCVILIQHQAQKPTNYLYFLSLAVSDLVLLGMLGMELEVEMRNRYPFLEGPVCYFKT 123
 DB 56 llvcmvivrngtkltpnylyflslavsdllvllgmlgmelevemhnyplflgpcyfk 115
 OY 124 ALFETVCFASILSTTVERVAAILHPFRALQSTRRALRILGIWGSVLSLEPNTS 183
 DB 116 alfetvcfasilsttvervvaailhpfraklqstrralrllglwgsvlsfslpnts 175
 OY 184 IHGKHFYFPNGSLVPSGATCTVYIKPMWYITFIOTSEFLYLLPMVITVLYLMAURL 243
 DB 176 ihgkfhfypngslvpsgactctvlykpmwlyitfioyseflyllpmtvlslylmrl 235
 OY 244 KKKSLEADGGMANIOPRCRSYVNMFLVLYFAICWAPFHIDRLEFSVEESLAA 303
 DB 236 kksleadeegmanioqrctsvnmflvlyfalcwaphidrlfseveeslaa 295
 OY 304 VFNLVHVSGVFYFLSSAVNPITYNLRSRFOAFQNVISFHKOMHSQHPOLPPAQRN 363
 DB 296 vfnlvhvsgvfyflssavnpilynlrsrfqafqnvlsfkhqmsqhdppqpqrn 355

OY 364 IFLETCHEVELTEDIGPOFCOSSMHNLSHPTA 396
 Db 356 iflechlmetledagpqfpgssinhnlta 388

RESULT 13

AAG63367
 ID AAG63367 standard; Protein: 395 AA.

XX AAG63367;

DT 15-OCT-2001 (first entry)

DE Amino acid sequence of a rat TGR-1 protein.

XX TGR-1; neuromedin U; hypertension; stress disease.

OS Rattus sp.

XX WO200157524-A1.

XX 09-AUG-2001.

XX 02-FEB-2001; 2001WO-JP00746.

XX 04-FEB-2000; 2000JP-0032773.

XX 24-FEB-2000; 2000JP-0052252.

XX 30-MAR-2000; 2000JP-0097886.

XX 19-JUN-2000; 2000JP-0187536.

XX (TAKE) TAKEDA CHEM IND LTD.

XX Hluma S, Shintani Y, Hosoya M, Fujii R, Moriya T, Matsui H;

XX Okubo S;

XX WPI; 2001-488917/53.

XX DR N-PSDB: AAA43078.

XX Identifying predicted or actual structures of two or more members of a

XX PT chemical or physical library by mass spectrometry comprising

XX PT correlating molecular mass measurements of two or more members with a

XX PT shared chemical history -

XX Claim 1; Page 89-91; 95pp; Japanese.

XX The present sequence represents a rat TGR-1 protein. The specification

XX CC describes a method of screening a compound, which is capable of binding

XX CC properties of neuromedin U to TGR-1. The method is useful for screening

XX CC preventative and remedies for hypertension, stress diseases, etc..

XX CC TGR-1 antagonists are also useful for treating the same diseases.

XX Sequence 395 AA;

XX Query Match 75.2%; Score 1643.5; DB 22; Length 395;

XX Best Local Similarity 78.9%; Pred. No. 1,1e-182;

XX Matches 310; Conservative 37; Mismatches 41; Indels 5; Gaps 1;

OY 4 MEKQNSMWTQOKLEDPFOKHNSTERYLAFLCGPRSHFLEPVSVVYPIFVNGVGN 63

Db 1 mglklnswlh----dpImkylnsteeYlahlcgprksdlsipysvayallflvngvn 55

OY 64 VLVLVLIVHOAMKTPNRYVLESLAVSOLVLLIGMPLEVEEMRNVPFLEGPVGCYFKT 123

Db 56 lllvmvavtrhqltkrptnyllfslavsdllvlllgmpleymhnyplfgpygcYkT 115

OY 124 ALFFETVCFASLITVTYSEVRYVALILHFRAKLQSTRRARILGIWGFVLSLPTMS 183

Db 116 alffetvcfaslsvtcsverYvalilHFRaklQSTRRARILGIWGFVLSLPTMS 175

OY 184 IHGIRFHYFPGSLVPGSAGATVIRPMWYFIQVTSFLFLLPMYIVSVLYLMLRL 243

Db 176 ihgikfhyfpgssavpgsaGATVIRpmwYFIQVTSFLFLLPMYIVSVLYLMLRL 235

OY 244 KKDKSLDEAGNANIORPCRSYKMLFVLYVFAICAPFHIDRLFFSFVEKSESIAA 303
 Db 236 krdesleankvaavnhiprsksvtkmlfvlvfaicwtpfhvdrllffsfveekeslaa 295

RESULT 14

AAB67806
 ID AAB67806 standard; Protein: 296 AA.

XX AAB67806;

DT 29-JUN-2001 (first entry)

DE Splice variant of G-protein coupled receptor IGS4A long version.

XX Human; G-protein coupled receptor; IGS4; IGS4A; IGS4B; schizophrenia;

XX nervous system disorder; psychiatric disorder; Parkinson's disease;

XX episodic paroxysmal anxiety; phobias; migraine; epilepsy; bulimia; stroke;

XX cardiovascular disease; heart failure; angina pectoris; obesity; emesis;

XX motility disorder; myocardial infarction; hypertension; dyslipidemia;

XX gastrointestinal disorder; inflammatory bowel disease; osteoporosis;

XX inflammation; infection; pain; cancer; immune disorder; allergy; sepsis;

XX gynecological disorder.

XX Homo sapiens.

XX WO200125269-A2.

XX 12-APR-2001.

XX 25-SEP-2000; 2000WO-EP09584.

XX 24-SEP-1999; 99EP-0203140.

XX 24-SEP-1999; 99NL-1013140.

XX 28-JUL-2000; 2000EP-0202683.

XX 31-JUL-2000; 2000US-0222047.

XX (SOLV) SOLVAY PHARM BV.

XX Deleersnijder W, Berger C, Loeken C, Nys G, Venema J;

XX WPI; 2001-273568/28.

XX DR N-PSDB: AAB60326.

XX New G-protein coupled receptors and the polynucleotides encoding them,

XX PT useful for preventing, ameliorating or correcting nervous system

XX PT disorders, cardiovascular diseases, dyslipidemias, inflammations, pain

XX PT or cancers -

XX Example 1b; Page 92-93; 102pp; English.

XX The present sequence represents a splice variant of the long version of a

XX human G-protein coupled receptor designated IGS4A. IGS4 exists in two

XX polymorphic forms, IGS4A and IGS4B. The IGS4 receptors and IGS4

XX polynucleotides are useful for preventing, ameliorating or correcting

XX dysfunctions or diseases. These diseases include peripheral nervous

XX system, psychiatric and central nervous system disorders

XX (e.g. schizophrenia, episodic paroxysmal anxiety, phobias, Parkinson's

XX disease, migraine, epilepsy, bulimia or stroke), cardiovascular diseases

XX (e.g. heart failure, angina pectoris, myocardial infarction or

XX hypertension), dyslipidemias, obesity, emesis, gastrointestinal

XX disorders (e.g. inflammatory bowel disease or motility disorders),

XX osteoporosis, inflammations, infections (e.g. bacterial, fungal,

XX protozoan or viral), pain, cancers, immune disorders, allergies,

CC sepsis or gynecological disorders. Agonists or antagonists of IG54 are
 CC effective with regard to disorders of the nervous system, including the
 CC central and peripheral nervous systems, disorders of the gastrointestinal
 CC system, cardiovascular system, skeletal muscle, thyroid, lung or
 CC genitourinary system, or immunological disease. The IG54 polynucleotides
 CC are useful as diagnostic reagents for detecting under-expression,
 CC overexpression or altered expression of IG54.

XX Sequence 296 AA:

Query Match 64.5%; Score 1409; DB 22; Length 296;
 Best Local Similarity 100.0%; Pred. No. 1.6e-155;
 Matches 270; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSGMEKLNASMTYQOKLEDPFOKHLNSTEYLAFLCGPRSHFPLPVSVVYPIFVGV 60
 DB 1 megmeklnasmtlyqokledepfokhlnsteylaflcgprrshfplpvsvvyplifvgy 60
 QY 61 IGVNLVCLVILQHQAMKPTNNYLFSLAVSDLLVLLGMPLEVEYEMWRNYPFLGPGVCY 120
 DB 61 IGVNLVCLVILQHQAMKPTNNYLFSLAVSDLLVLLGMPLEVEYEMWRNYPFLGPGVCY 120
 QY 121 FRTALFETVCFASISITTVSVRVAIIHPFRAKQSTRRRRLRLIGTWGCVLFSLP 180
 DB 121 FRTALFETVCFASISITTVSVRVAIIHPFRAKQSTRRRRLRLIGTWGCVLFSLP 180
 QY 181 NMSIHGIRKHYFPNGSLVPGSATCTYIKPMWYNFIQVTSFLFYLLPMTVISLVLYLMA 240
 DB 181 nsihgirkyfngslvpgsatctyikpmwynfiqvtsflfyllpmtvislvlylma 240
 QY 241 LRLKDKSLEADGNANIORPCRKSVNKM 270
 DB 241 lrlkdkslleadgnaniorpcrksvnkm 270

RESULT 15

ID AAB67807 standard; Protein: 293 AA.

AC AAB67807;

DT 29-JUN-2001 (first entry)

DE Splice variant of G-protein coupled receptor IG54A short version.

XX Human; G-protein coupled receptor; IG54; IG54A; IG54B; schizophrenia;
 KW nervous system disorder; psychiatric disorder; Parkinson's disease;
 KW episodic paroxysmal anxiety; phobia; migraine; epilepsy; bulimia; stroke;
 KW cardiovascular disease; heart failure; angina pectoris; obesity; emesis;
 KW motility disorder; myocardial infarction; hypertension; dyslipidemia;
 KW gastrointestinal disorder; inflammatory bowel disease; osteoporosis;
 KW inflammation; infection; pain; cancer; immune disorder; allergy; sepsis;
 KW gynecological disorder.

OS Homo sapiens.

PN WO200125269-A2.

PD 12-APR-2001.

PF 25-SEP-2000; 2000WO-EP09584.

XX 24-SEP-1999; 99EP-0203140.

PR 24-SEP-1999; 99NL-1013140.

PR 28-JUL-2000; 2000EP-0202683.

PR 31-JUL-2000; 2000US-0222047.

XX (SOLV) SOLVAY PHARM BV.

XX Delecrn1jder W, Berger C, Loeken C, Nys G, Venema J;
 DR WPI; 2001-273568/28.

DR N-PSDB; AAF80327.
 XX New G-protein coupled receptors and the polynucleotides encoding them,
 PT useful for preventing, ameliorating or correcting nervous system
 PT disorders, cardiovascular diseases, dyslipidemias, inflammations, pain
 PT or cancers

XX Example 1b; Page 96-97; 102pp; English.

XX The present sequence represents a splice variant of the short version of
 CC a human G-protein coupled receptor designated IG54A. IG54 exists in two
 CC polymorphic forms, IG54A and IG54B. The IG54 receptors and IG54
 CC polynucleotides are useful for preventing, ameliorating or correcting
 CC dysfunction or diseases. These diseases include peripheral nervous
 CC system, psychiatric and central nervous system disorders
 CC (e.g. schizophrenia, episodic paroxysmal anxiety, phobia, Parkinson's
 CC disease, migraine, epilepsy, bulimia or stroke), cardiovascular diseases
 CC (e.g. heart failure, angina pectoris, myocardial infarction or
 CC hypertension), dyslipidemias, obesity, emesis, gastrointestinal
 CC disorders (e.g. inflammatory bowel disease or motility disorders),
 CC osteoporosis, inflammations, infections (e.g. bacterial, fungal,
 CC protozoan or viral), pain, cancers, immune disorders, allergies,
 CC sepsis or gynecological disorders. Agonists or antagonists of IG54 are
 CC effective with regard to disorders of the nervous system, including the
 CC central and peripheral nervous systems, disorders of the gastrointestinal
 CC system, cardiovascular system, skeletal muscle, thyroid, lung or
 CC genitourinary system, or immunological disease. The IG54 polynucleotides
 CC are useful as diagnostic reagents for detecting under-expression,
 CC overexpression or altered expression of IG54.

XX Sequence 293 AA:

Query Match 63.8%; Score 1394; DB 22; Length 293;
 Best Local Similarity 100.0%; Pred. No. 8.8e-154;
 Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 MEKLNASMTYQOKLEDPFOKHLNSTEYLAFLCGPRSHFPLPVSVVYPIFVGVGN 63
 DB 1 meklngasmtlyqokledepfokhlnsteylaflcgprrshfplpvsvvyplifvgygn 60
 QY 64 VLVCLVILQHQAMKPTNNYLFSLAVSDLLVLLGMPLEVEYEMWRNYPFLGPGVCYFKT 123
 DB 64 VLVCLVILQHQAMKPTNNYLFSLAVSDLLVLLGMPLEVEYEMWRNYPFLGPGVCYFKT 120
 QY 124 ALFETVCFASISITTVSVRVAIIHPFRAKQSTRRRRLRLIGTWGCVLFSLPNTS 183
 DB 124 alfetvcfasisiltvsvrvaiihpfrakqstrrrrlrligtwgcvlfslpnts 180
 QY 184 IHGIRKHYFPNGSLVPGSATCTYIKPMWYNFIQVTSFLFYLLPMTVISLVLYLMA 243
 DB 184 ihgirkhyfngslvpgsatctyikpmwynfiqvtsflfyllpmtvislvlylma 240
 QY 244 KDKSLEADGNANIORPCRKSVNKM 270
 DB 244 kdkslleadgnaniorpcrksvnkm 267

Search completed: April 22, 2002, 14:13:39
 Job time: 52 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2002, 14:19:41 ; Search time 17.05 Seconds
(without alignments)
849.420 Million cell updates/sec

Title: US-09-609-146-25
Perfect score: 2076
Sequence: 1 MCKLENASWIDPLMKYLNLS.....GOSSIHNTLTTPACAGEVP 395

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Swissprot_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	513	24.7	424	1	NTR1_RAT
2	508.5	24.5	424	1	NTR1_MOUSE
3	499	24.0	418	1	NTR1_HUMAN
4	477.5	23.0	412	1	GP38_HUMAN
5	465.5	22.4	364	1	GHSR_RAT
6	465	22.4	366	1	GHSR_PIG
7	463.5	22.3	366	1	GHSR_HUMAN
8	437	21.1	410	1	NTR2_HUMAN
9	430.5	20.7	416	1	NTR2_MOUSE
10	426.5	20.5	417	1	NTR2_MOUSE
11	416	20.0	398	1	TRFR_HUMAN
12	414	19.9	398	1	TRFR_BOVIN
13	413.5	19.9	395	1	TRFR_CHICK
14	412	19.8	393	1	TRFR_MOUSE
15	412	19.8	398	1	TRFR_SHEEP
16	409	19.7	402	1	NK2R_CAVPO
17	403.5	19.4	412	1	TRFR_RAT
18	400.5	19.3	400	1	OPRM_HUMAN
19	399.5	19.2	384	1	NK2R_BOVIN
20	393.5	19.0	401	1	OPRM_BOVIN
21	393	18.9	398	1	NK2R_HUMAN
22	387.5	18.7	385	1	NK3R_MOUSE
23	387	18.6	398	1	OPRM_MOUSE
24	380.5	18.3	452	1	NK3R_RAT
25	380	18.3	398	1	OPRM_RAT
26	378.5	18.2	401	1	OPRM_PIG
27	375.5	18.1	384	1	NK2R_RABIT
28	373	18.0	384	1	NK2R_MOUSE
29	372	17.9	408	1	NK1R_RANCA
30	370.5	17.8	467	1	NK3R_RABIT
31	368	17.7	390	1	NK2R_RAT
32	366.5	17.7	465	1	NK3R_HUMAN
33	363.5	17.5	384	1	SSR4_RAT

34	362.5	17.5	363	1	SSR5_RAT	P30938	rattus norv
35	362	17.4	380	1	OPRK_RAT	P34975	rattus norv
36	362	17.4	384	1	NK2R_MESAU	P51144	mesocricetu
37	362	17.4	453	1	GP39_HUMAN	O43194	homo sapien
38	361	17.4	504	1	TLR1_DROME	P30974	drosophila
39	360.5	17.4	362	1	SSR5_MOUSE	O08858	mus muscullu
40	360	17.3	380	1	OPRK_CAVPO	P70031	xenopus lae
41	360	17.3	453	1	CCRR_XENLA	P41144	cavia porce
42	359.5	17.3	380	1	OPRK_HUMAN	P41145	homo sapien
43	359	17.3	369	1	SSR2_HUMAN	P30874	homo sapien
44	359	17.3	440	1	NK4R_HUMAN	P30098	homo sapien
45	358	17.2	368	1	SSR2_BOVIN	P34993	bos taurus

ALIGNMENTS

RESULT	ID	STANDARD	PRT	424 AA.
1	NTR1_RAT			
AC	P20789:			
DT	01-FEB-1991 (rel. 17, last sequence update)			
DT	01-FEB-1991 (rel. 17, last sequence update)			
DE	NEUROKININ RECEPTOR TYPE 1 (NT-R-1) (HIGH-AFFINITY LEVOCABASTINE-SENSITIVE NEUROKININ RECEPTOR) (NTRH).			
GN	NTR1.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Brain;			
RX	MEDLINE=90297956; PubMed=1694443;			
RA	Tanaka K., Masu M., Nakanishi S.;			
RT	*Structure and functional expression of the cloned rat neurotensin receptor.*			
RL	Neuron 4:847-854(1990).			
CC	-1- FUNCTION: RECEPTOR FOR THE TRIDECAPEPTIDE NEUROKININ. IT IS ASSOCIATED WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-3-OH PHOSPHOLIPASE C (PLC) AND A CALCIUM SECOND MESSENGER SYSTEM.			
CC	-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.			
CC	-1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.			
CC	HIGHEST TO TACHYKININS RECEPTORS.			
DR	PIR: JH0164; JH0164.			
DR	GCRD: GCR_0219; -			
DR	InterPro: IPR000276; GPCR_Rhodpsn.			
DR	Pfam: PF00001; 7tm_1; 1.			
DR	PRINTS: PR00237; GPCR_Rhodopsin.			
DR	PROSITE: PS00237; G-PROTEIN RECEPTOR FL-1; 1.			
DR	PROSITE: PS00262; G-PROTEIN RECEPTOR FL-2; 1.			
KW	G-protein coupled receptor; Transmembrane; Glycoprotein;			
KW	phosphorylation; Lipoprotein; Palmitate.			
FT	DOMAIN 1	64		
FT	TRANSMEM 65	87		
FT	DOMAIN 88	96		
FT	TRANSMEM 97	121		
FT	DOMAIN 122	143		
FT	TRANSMEM 144	165		
FT	DOMAIN 166	188		
FT	TRANSMEM 189	210		
FT	DOMAIN 211	235		
FT	TRANSMEM 236	260		
FT	DOMAIN 261	308		
FT	TRANSMEM 309	330		
FT	DOMAIN 331	348		
FT	TRANSMEM 349	372		
FT	DOMAIN 373	424		
FT	CARBOHYD 4			
FT	CARBOHYD 38			
FT	CARBOHYD 42			
FT	DISULFID 142	225		

FT LIPID 388 388 PALMITATE (POTENTIAL).
 SQ SEQUENCE 424 AA: 47054 MW: A9C2F7EAF8D9BCD3 CRC64;
 Query Match 24.7%; Score 513; DB 1; Length 424;
 Best Local Similarity 31.6%; Pred. No. 2,6e-26;
 Matches 119; Conservative 76; Mismatches 113; Indels 68; Gaps 13;

QY 19 NSTREYLAHLGCPKRSLSLP-----VSVAVALFLVGVGNLVCVAVIHR--QT 67
 DB 42 NTSESDA---GP NSLDVNTDIYSKVLTATITLAFVGVGNSTATTIARKKSLSQ 97
 QY 68 LKPTPTNYLFLSLAVSDLLVLLGMPLEIYE-MMHNPFLGPGVC--YFKTALFETVCF 123
 DB 98 LQSTVHYHGLSLASDLLILLAMPVELYNFIWHHPWARGDAGCGRGYF--LRDCTY 154
 QY 124 ASLSTVTVVERVVALVHFRPAKLESTRRALRILSLVSVSVSFLPMTSHGKIFQH 183
 DB 155 ATALNVAISLSEVERLACHPFAKTLMSRSRTKFKISAILASLALPMLFTMGL--QN 212
 QY 184 FPNQSSVPGSATCTVTKPMVNYLIIOATSEFLFYILPMTLISVLYLMGLRLKRDESLA 243
 DB 213 RSGDGTIPGGLVCTPIYDTATVAVIOVNTFMSPFLPMLVLSIL-----NTVIA 261
 QY 244 NKVAVINIR-----PSRKSVTK---MLFVLVLAFCW 273
 DB 262 NKLTVVHQAEQGRVCTVGHNGLEHSTFNMTIEPGVQALRHGVLLRAVIAFVVCW 321
 QY 274 TPPIHVRLFFSFV--EEMTESLAVENLIVSGVFYLSAVNPITLYNLSRRFAFR 331
 DB 322 LPIYHRLMFCYISDEQMTFLFDYHYFYMLNALFYSSAINPILYNLSANFRQVFL 381
 QY 332 NVVSPCK-WCHPRHR 346
 DB 382 STLACLCPGMRHRKK 397

RESULT 2
 NTNL_MOUSE
 ID NTNL_MOUSE STANDARD; PRT: 424 AA.
 AC 088319:
 DT 20-AUG-2001 (Rel. 40, Created)
 DT 20-AUG-2001 (Rel. 40, Last sequence update)
 DE 20-AUG-2001 (Rel. 40, Last annotation update)
 DE NEUTROSENSIN RECEPTOR TYPE 1 (NT-R-1).
 GN NTSR1 OR NTSR.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 ON NCBI_Taxid=10090;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE-Brain:
 RA Snider J., Sano H., Ohta M.;
 RT "Neurotensin receptor type 1,"
 RT Submitted (Aug-1998) to the EMBL/GenBank/DBJ databases.
 CC - FUNCTION: RECEPTOR FOR THE TRIDECAPEPTIDE NEUROTENSIN. IT IS
 CC ASSOCIATED WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-
 CC CALCIUM SECOND MESSENGER SYSTEM.
 CC - SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC - SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC - SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC HIGHEST TO TACHYKININS RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: AB017027; BAA33013.1; -
 DR MGD: MGI:97386; Ntsr.

DR InterPro: IPR000276; GPCR_Rhodopsin.
 DR Pfam: PF00001; 7tm.1; 1.
 DR PRINTS: PR000237; GPCRHOPOPSN.
 DR PROSITE: PS00237; G_PROTEIN_RECP_F1_1; 1.
 DR PROSITE: PS50262; G_PROTEIN_RECP_F1_2; 1.
 KW G-protein coupled receptor; Transmembrane; glycoprotein;
 KW Phosphorylation; Lipoprotein; Palmitate.
 FT DOMAIN 1 64
 FT TRANSMEM 65 87
 FT DOMAIN 88 96
 FT TRANSMEM 97 121
 FT DOMAIN 122 143
 FT TRANSMEM 144 165
 FT DOMAIN 166 188
 FT TRANSMEM 189 210
 FT DOMAIN 211 235
 FT TRANSMEM 236 260
 FT DOMAIN 261 308
 FT TRANSMEM 309 330
 FT DOMAIN 331 348
 FT TRANSMEM 349 372
 FT DOMAIN 373 424
 FT CARBOHYD 4 4
 FT CARBOHYD 38 38
 FT CARBOHYD 42 42
 FT CARBOHYD 211 211
 FT DISULFID 141 224
 FT LIPID 388 388
 SQ SEQUENCE 424 AA: 47216 MW: 8E9A723171A48711 CRC64;
 PALMITATE (POTENTIAL).

Query Match 24.5%; Score 508.5; DB 1; Length 424;
 Best Local Similarity 32.2%; Pred. No. 5,1e-26;
 Matches 116; Conservative 70; Mismatches 99; Indels 75; Gaps 13;

QY 40 VSAVALFLVGVGNLVCVAVIHR--QTLKPTNYLFLSLAVSDLLVLLGMPLEIY 96
 DB 66 VTAVALFLVGVGNSTATTIARKKSLSQSLQSTVHYHGLSLASDLLILLAMPVELY 125
 QY 97 E-MMHNPFLGPGVC--YFKTALFETVCFASLSTVTVVERVVALVHFRPAKLESTR 152
 DB 126 NFIWVHHHPMAFGDAGCGRYF--LRDCTYVAAALNVAISLSEVERLACHPFAKTLMSR 182
 QY 153 RRALRILSLVSVSVSFLPMTSHGKIFQHPPNCSVPGSATCTVTKPMV 204
 DB 183 SRTRKFTSAITWLSALLAVMLFTMGIONRSAG--QH-----PGGLVCTPIYDTAT 232
 QY 205 YNLIOATSEFLFYILPMTLISVLYLMGLRLKRDESLANKVAVNIHR----- 252
 DB 233 VKVVIQVNTFMSPFLPMLVLSIL-----NTVIAKLTVMVHQAEQGRVCTVG 281
 QY 253 -----PSRKSVTK---MLFVLVLAFCWTPPHVDRLPFSFV--EEMTE 291
 DB 282 THNSLEHSTFNMSEIPGRVQALRHGVLLRAVIAFVVCCLPYHRLMFCYISDQMTT 341
 QY 292 SLAAVFNLIHVGVEFYLSAVNPITLYNLSRRFAFRNVVSPCK-WCHPRHRPGCF 350
 DB 342 FLDFDFHYEFMLNALFYSSAINPILYNLSANFRQVFLSTLACLCPGM--RRRRKKRP 399

RESULT 3
 NTNL_HUMAN
 ID NTNL_HUMAN STANDARD; PRT: 418 AA.
 AC P30989:
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE NEUTROSENSIN RECEPTOR TYPE 1 (NT-R-1) (HIGH-AFFINITY LEVOCABASTINE-
 DE INSENSITIVE NEUROSENSIN RECEPTOR) (NTRH).
 GN NTSR1 OR NTSR.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93154505; PubMed=8381365;
 RA Vita N., Laurent P., Lefort S., Chalou P., Dumont X., Kaghad M.,
 RA Gully D., Le Fur G., Ferrara P., Caput D.;
 RT Cloning and expression of a complementary DNA encoding a high
 RT affinity human neurotensin receptor.;
 RL FEBS Lett. 317:139-142(1993).
 CC -1- FUNCTION: RECEPTOR FOR THE TRIDECAPEPTIDE NEUROTENSIN. IT IS
 CC ASSOCIATED WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-
 CC CALCIUM SECOND MESSENGER SYSTEM.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC HIGHEST TO TCHYKININS RECEPTORS.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 CC EMBL: X70070; CAA49675.1; -
 CC PIR: S29506; S29506.
 CC GCRDB: GCR_0577; -
 CC GCRDB: GCR_2067; -
 CC MIM: 162651; -
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR Pfam: PF00001; 7tm_1; 1.
 DR PROSITE: PS00237; G-PROTEIN_RECEP_FL1; 1.
 DR PROSITE: PS50262; G-PROTEIN_RECEP_FL2; 1.
 KM G-protein coupled receptor; Transmembrane; Glycoprotein;
 KM Phosphorylation; Lipoprotein; Palmitate.
 FT DOMAIN 1 63
 FT DOMAIN 64 86
 FT DOMAIN 87 95
 FT TRANSMEM 96 120
 FT DOMAIN 121 142
 FT TRANSMEM 143 164
 FT DOMAIN 165 187
 FT TRANSMEM 188 209
 FT DOMAIN 210 234
 FT TRANSMEM 235 259
 FT DOMAIN 260 303
 FT TRANSMEM 304 325
 FT DOMAIN 326 343
 FT TRANSMEM 344 367
 FT DOMAIN 368 418
 FT CARBOHYD 4 4
 FT CARBOHYD 37 37
 FT CARBOHYD 41 41
 FT DISULFID 141 224
 FT LIPID 383 383
 SQ SEQUENCE 418 AA: 46288 MW: BBBDIIEC2BE6E390 CRC64:

 Query Match 24.0%; Score 499; DB 1: Length 418;
 Best Local Similarity 33.0%; Pred. No. 2e-25;
 Matches 116; Conservative 65; Mismatches 103; Indels 68; Gaps 12;

OY 205 YNLIQATSPFLYILPMTLISLYLMLGLRKDESLKAVNINR----- 252
 DB 233 VKVIVQNTFMSTFPMVAVISL-----NTIIANKLIVNVRQAEOGVCTVGS 281
 OY 253 -----PSR-----KSTYKMLFVLYLVAICWTPFHDRLFSEFV--EEMTESLAIV 296
 DB 282 EHSFMSAIEPGRVQALRHGVRLRAVIAVFAVCWMLPYHVRIMPCYISDEQWTPFLYDP 341
 OY 297 ENLIHYVSGFFFLSSAVNPITINLSRRFRANVSTCK-WCHPRHRP 347
 DB 342 YHVEYVNTNALFVYSTINPILNLVANSRHFLETLACLCPVWRRRRKRP 393

 RESULT 4
 GP38_HUMAN STANDARD: PRT: 412 AA.
 ID GP38_HUMAN
 AC 043193;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 20-APR-2001 (Rel. 40, Last annotation update)
 DE PUTATIVE G-PROTEIN-COUPLED RECEPTOR GPR38.
 GN GPR38.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98110578; PubMed=9441746;
 RA McKee K.K., Tan C.P., Palyna O.C., Liu J., Feighner S.D.,
 RA Hreniuk D.L., Smith R.G., Howard A.D., van der Ploeg L.H.T.;
 RT Cloning and characterization of two human G-protein-coupled receptor
 RT genes (GPR38 and GPR39) related to the growth hormone secretagogue
 RT and neurotensin receptors.;
 RL Genomics 46:426-434(1997).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- TISSUE SPECIFICITY: EXPRESSED ONLY IN THYROID, STOMACH, AND BONE
 CC MARROW.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 CC EMBL: AF034632; AAC26081.1; -
 CC GCRDB: GCR_2494; -
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR Pfam: PF00001; 7tm_1; 2.
 DR PRINTS: PR00237; GPCR_Rhodopsn.
 DR PROSITE: PS00237; G-PROTEIN_RECEP_FL1; 1.
 DR PROSITE: PS50262; G-PROTEIN_RECEP_FL2; 1.
 KM G-protein coupled receptor; Transmembrane; Glycoprotein;
 KM Extracellular; (POTENTIAL).
 FT DOMAIN 1 35
 FT TRANSMEM 36 56
 FT DOMAIN 57 74
 FT TRANSMEM 75 94
 FT DOMAIN 95 112
 FT TRANSMEM 113 134
 FT DOMAIN 135 157
 FT TRANSMEM 158 178
 FT DOMAIN 179 246
 FT TRANSMEM 247 270
 FT DOMAIN 271 298
 FT TRANSMEM 299 320
 FT DOMAIN 321 334
 FT TRANSMEM 335 358
 FT DOMAIN 359 412

 Query Match 24.0%; Score 499; DB 1: Length 418;
 Best Local Similarity 33.0%; Pred. No. 2e-25;
 Matches 116; Conservative 65; Mismatches 103; Indels 68; Gaps 12;

FT DISULFID 111 235 BY SIMILARITY.
 FT CARBOHYD 6 6 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 192 192 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 412 AA: 45344 MM: C13FF6165012DEF3 CRC64;

Query Match 23.0%; Score 477.5; DB 1; Length 412;
 Best Local Similarity 30.8%; Pred. No. 4.8e-24;
 Matches 112; Conservative 81; Mismatches 114; Indels 57; Gaps 9;

QY 29 CGKRSLSLPVSVAVALIFLVGMGLVCMYVHRQTLKTPNTYLSLAVSDLLVL 88
 DB 30 CSFPLPALVAVAVCLFPLVGVSGVNVVTMLGRYDRKRTNTLVLGMAVSDLLI-L 88
 QY 89 LGMPLLEYEMHNPFLFPGVGCYKALFETVCFASILSTVSVRYAVIAPRAKL 148
 DB 89 LGLPFDLYRMRKRPWFGLCLRLSLYVGGCYATLIMHTALSVRYIALICPLRARV 148
 QY 149 ESTRRRLRLILSVSPSVFSLPNTSHIGIKFQHPNGSSVPG-----SATCTVTK 200
 DB 149 LVTRRRYRALIAYLMAVALLSAGPFLVGV--EQDPGISVPELNGTARIASSPLASP 206
 QY 201 PMWV-----YLLIQAISFLFIYLPMLISV 226
 DB 207 PLMLSRAPPSPPSGPETAAMALFSRECRSPAQLGALRYMLMVTAYEFF-LPFLCLSI 265
 QY 227 LVYLMGLRLKRDSELEANKYAVNHRPSRKSVTKMLFVLVLFVAFICWTPHYDRLFPSV 286
 DB 266 LVYGLIGLELSSRRPLRGPAASGREHROT-V-RVLLVVLAFITCLPHVGHIIITYNT 324
 QY 287 EEMTESLAAVFNLIHVSQVFFYLSAVNPITYNLSRRRR-AAFRNVSPTECKMCHPRH 345
 DB 325 ED--SRMWFYSQYFNIVALQLFYLASINPILYLSLKRYRAAFKLLLA-----RKS 375
 QY 346 RPOG 349
 DB 376 RPKG 379

RESULT 5

GHSR_RAT STANDARD; PRT; 364 AA.

AC 008725;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE GROWTH HORMONE SECRETAGOCUE RECEPTOR TYPE 1 (GHS-R) (GH-RELEASING PEPTIDE RECEPTOR) (GHRP).
 GN GHSR.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Rattus.
 NC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;

RN 11
 RP SEQUENCE FROM N.A.
 RC TISSUE-Pituitary;
 RX MEDLINE=97246555; PubMed=9092793;
 RA McKee K.K., Palyha O.C., Feighner S.D., Hreniuk D.L., Tan C., Phillips M.S., Smith R.G., der Ploeg L.H.T., Howard A.D., "Molecular analysis of rat pituitary and hypothalamic growth hormone secretagogue receptors."
 RT secretagogue receptors."
 RL MOL. Endocrinol. 11:415-423(1997).
 RN 12
 RP SEQUENCE OF 1-240 FROM N.A.
 RC STRAIN-Wistar; Tissue-Pituitary;
 RX MEDLINE=98100386; PubMed=9437732;
 RA Yokoe R., Sato M., Matsubara S., Ohye H., Mimi M., Murao K., Takahara J., "Molecular cloning and gene expression of growth hormone-releasing peptide receptor in rat tissues."
 RT peptide receptor in rat tissues."
 RL Peptides 19:15-20(1998).

CC -1- FUNCTION: RECEPTOR FOR GROWTH HORMONE RELEASING PEPTIDES (GHRP) AS WELL AS NON-PEPTIDE, LOW MOLECULAR WEIGHT SECRETAGOGUES (E.G. L-

CC 692,429, MK-0677). THIS RECEPTOR IS COUPLED TO G-ALPHA-11
 CC PROTEINS (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: U94321; AAC5156.1; -
 CC EMBL: AB001982; BAA2177.1; ALI_INIT.
 CC GCRDB: GCR_1383; -
 CC InterPro: IPR003905; GHS1_receptor.
 CC InterPro: IPR000276; GPCR_Rhodopsin.
 CC Pfam: PF00001; 7tm_1; 1.
 CC PRINTS: PR00237; GPCR_Rhodopsin.
 CC PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; 1.
 CC PROSITE: PS0262; G_PROTEIN_RECEP_F1_2; 1.
 CC K6 G-protein coupled receptor; Transmembrane; Glycoprotein.
 CC FT DOMAIN 1 40
 CC FT TRANSMEM 41 66
 CC FT DOMAIN 67 72
 CC FT TRANSMEM 73 96
 CC FT DOMAIN 97 117
 CC FT TRANSMEM 118 139
 CC FT DOMAIN 140 162
 CC FT TRANSMEM 163 183
 CC FT DOMAIN 184 211
 CC FT TRANSMEM 212 235
 CC FT DOMAIN 236 263
 CC FT TRANSMEM 264 285
 CC FT DOMAIN 286 302
 CC FT TRANSMEM 303 326
 CC FT DOMAIN 327 364
 CC FT DISULFID 115 197
 CC FT CARBOHYD 13 13
 CC FT CARBOHYD 26 26
 CC SQ SEQUENCE 364 AA: 40963 MM: DCBF559BE061EE9 CRC64;

Query Match 22.4%; Score 465.5; DB 1; Length 364;
 Best Local Similarity 34.6%; Pred. No. 2.5e-23;
 Matches 104; Conservative 65; Mismatches 107; Indels 25; Gaps 6;

QY 40 VSAVALIFLVGMGLVCMYVHRQTLKTPNTYLSLAVSDLLVLGMPLEYEMW 99
 DB 45 VTAACVALFVGVISGNLLTMLVSRFRELRTTNLVLSSMAFSLFLIC-MPLDLVRLM 103
 QY 100 HNPFLFPGVGCYKALFETVCFASILSTVSVRYAVIAPRAKLESTRRALRL 159
 DB 104 QYRPMNGDGLCKLDFQVSSCYATVLTITLALSVRYFAICPLRAKVVTGRVNLVI 163
 QY 160 SLVWFSFVSFLPNTSHIGIKFQHPNGSSVPGSATCTVTK-----PMWYVLLI 209
 DB 164 LVITAAVAFCSAGPLFVLVGVDEH---NGTDPDPTNEGRATEFAVRSGLLTVMWV----- 215
 QY 210 QATSEFLFIYLPMLISVLYLMLGLRLKRDSELEANKYAVNHRPSRKSVTKMLFVLVLF 269
 DB 216 ---SVVEFLPVLCFLVLYSLIGRKLRRRGDA--VGASLRDQNHQYTKMLAVVFAF 270
 QY 270 AICWTPHYDRLFPS-FVEEMTESLAAVFNLIHVSQVFFYLSAVNPITYNLSRRFRA 328
 DB 271 ILCWLPHYVGRYLSKSEFEGSLIAIOISQYCNLVSEFLVFLSAINPILYINIMSKRYV 330
 QY 329 A 329
 DB 331 A 331

```

RESULT 6
GHSR_PIG STANDARD: PRT: 366 AA.
AC 095234: 095235:
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE GROWTH HORMONE SECRETAGOGUE RECEPTOR TYPE 1 (GHS-R) (GH-RELEASING
PEPTIDE RECEPTOR) (GHRP).
GN GHSR.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
OX NCBI_TaxID=9823:
RN
RP SEQUENCE FROM N.A.
RC STRAIN=YORKSHIRE; TISSUE=Plutitary;
RX MEDLINE=96337998; Pubmed=8688086;
RA Howard A.D., Feighner S.D., Cully D.F., Arena J.P.,
RA Liberator P.A., Rosenblum C.I., Hamelin M., Hrenluk D.L.,
RA Palaya O.C., Anderson J., Pares P.S., Diaz C., Chou M., Liu K.K.,
RA McKee K.R., Pong S.-S., Chung L.-Y., Elbrecht A., Dashkevitz M.,
RA Heavens R., Rigby M., Srinathsinghji D.J.S., Dean D.C., Mellillo D.G.,
RA Patchett A.A., Nargund R., Griffin P.R., Demartino J.A., Gupta S.K.,
RA Schaeffer J.M., Smith R.G., van der Ploeg L.H.T.;
RT "A receptor in pituitary and hypothalamus that functions in growth
hormone release."
RL Science 273:974-977(1996).
CC -1- FUNCTION: RECEPTOR FOR GROWTH HORMONE RELEASING PEPTIDES (GHRP) AS
WELL AS NON-PEPTIDE, LOW MOLECULAR WEIGHT SECRETAGOGUES (E.G. L-
692,429, MK-0677). THIS RECEPTOR IS COUPLED TO G-ALPHA-11
CC CC PROTEINS. BINDING POTENCY FOR THE SECRETAGOGUES IS IN THE ORDER:
MK-0677 > GHRP-2 > GHRP-6
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
-1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1A (SHOWN HERE) AND 1B; ARE
PRODUCED BY ALTERNATIVE SPLICING. ISOFORM 1B APPEARS NOT TO BIND
SECRETAGOGUES.
CC -1- TISSUE SPECIFICITY: PITUITARY AND HYPOTHALAMUS.
-1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
-----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC EMBL: U60178; AAC48630.1;
DR EMBL: U60180; AAC48631.1;
DR GCRDB: GCR1533;
DR GCRDB: GCR1539;
DR InterPro: IPR003905; GHS1_receptor.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PR00237; GPCR_RHODOPSIN.
DR PROSITE: PS00237; G-PROTEIN_RECEP_FL_1; 1.
DR PROSITE: PS00262; G-PROTEIN_RECEP_FL_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
Alternative splicing.
FT DOMAIN 1 40 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 41 66 1 (POTENTIAL).
FT DOMAIN 67 72 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 73 96 2 (POTENTIAL).
FT DOMAIN 97 117 3 (POTENTIAL).
FT TRANSSEM 118 139 4 (POTENTIAL).
FT DOMAIN 140 162 5 (POTENTIAL).
FT TRANSSEM 163 183 6 (POTENTIAL).
FT DOMAIN 184 211 7 (POTENTIAL).
FT TRANSSEM 212 235 8 (POTENTIAL).
FT DOMAIN 236 263 9 (POTENTIAL).
FT TRANSSEM 264 285 10 (POTENTIAL).
FT DOMAIN 286 302 11 (POTENTIAL).

```

```

FT TRANSSEM 303 326 7 (POTENTIAL).
FT DOMAIN 327 366 CYTOPLASMIC (POTENTIAL).
FT DISULFID 116 198 BY SIMILARITY.
FT CARBOHYD 13 13 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 27 27 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 266 289 AVVFAPFLCPLPVPVGRVLRFSK -> GGSOCALFELSLPG
FT VARSPLIC 290 366 PLHSSCLEFSSP (IN ISOFORM 1B).
FT VARSPLIC 366 411 MISSING (IN ISOFORM 1B).
SQ SEQUENCE 366 AA: 41194 MW: 268503EF61B7C1C CAC64:

Query Match 22.4%; Score 465; DB 1; Length 366;
Best Local Similarity 34.2%; Pred. No. 2,7e-23;
Matches 103; Conservative 68; Mismatches 106; Indels 24; Gaps 6;

OY 40 VSAVALIFLVGNGNLLVCVYVHQTLPPTNYFLSLAVSDLLVLLGPLEIYEMW 99
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 46 VTATCVALFVVGAGNLLMTLVSVSRFRMTTNYLSMAFSDDLPLFC-MPLDLFLM 104
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 100 HNYPLPFGVCCYFETALTFFVNCFSILSVTVSVSRVAVIHPRAKLESTRRALRL 159
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 105 QYRPNNLNLCKLRFQFVSESCYTATVLTALSVRYFALCFPLRAVVVTKGVKLV 164
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 160 SLVMSFSVYFSLPNTSINGIKRFQHPNCSVPGSATCTVTK-----PMVYVNDII 209
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 165 LVYMAVACSGAPFIVLVGVEHD---NGTDPRDTECKATEAVAVNSGLITVWVWV----- 216
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 210 QATSFLEYILPMTLSVLYLMLGRLKDESLKAVANVNIHRSKSVTKMLFVLVLF 269
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 217 --SSVFELPVEFCLTVLYSLGRKLMRRKREA-AVGSLLDQNHKTKMLAVVPAF 272
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 270 AICWPFVVDRLFEF-FVEEMTESLAAVFNLIHVAGFVYLSAVNPIIYNLSRRRA 328
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 273 ILCWLPFVHGRILSKSLPGSVETAQISQYCNLVSEVFLYLAALNPLINIMSKYRV 332
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 329 A 329
DB 333 A 333

RESULT 7
GHSR_HUMAN STANDARD: PRT: 366 AA.
AC 092847: 092848:
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE GROWTH HORMONE SECRETAGOGUE RECEPTOR TYPE 1 (GHS-R) (GH-RELEASING
PEPTIDE RECEPTOR) (GHRP).
GN GHSR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606:
RN
RP SEQUENCE FROM N.A.
RC TISSUE=Plutitary;
RX MEDLINE=96337998; Pubmed=8688086;
RA Howard A.D., Feighner S.D., Cully D.F., Arena J.P.,
RA Liberator P.A., Rosenblum C.I., Hamelin M., Hrenluk D.L.,
RA Palaya O.C., Anderson J., Pares P.S., Diaz C., Chou M., Liu K.K.,
RA McKee K.R., Pong S.-S., Chung L.-Y., Elbrecht A., Dashkevitz M.,
RA Heavens R., Rigby M., Srinathsinghji D.J.S., Dean D.C., Mellillo D.G.,
RA Patchett A.A., Nargund R., Griffin P.R., Demartino J.A., Gupta S.K.,
RA Schaeffer J.M., Smith R.G., van der Ploeg L.H.T.;
RT "A receptor in pituitary and hypothalamus that functions in growth
hormone release."
RL Science 273:974-977(1996).
CC -1- FUNCTION: RECEPTOR FOR GROWTH HORMONE RELEASING PEPTIDES (GHRP) AS
WELL AS NON-PEPTIDE, LOW MOLECULAR WEIGHT SECRETAGOGUES (E.G. L-
692,429, MK-0677). THIS RECEPTOR IS COUPLED TO G-ALPHA-11
CC CC PROTEINS. BINDING POTENCY FOR THE SECRETAGOGUES IS IN THE ORDER:
MK-0677 > GHRP-2 > GHRP-6.

```

CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1A (SHOWN HERE) AND 1B; ARE
 CC PRODUCED BY ALTERNATIVE SPLICING. ISOFORM 1B APPEARS NOT TO BIND
 CC SECRETAGOGUES.
 CC -1- TISSUE SPECIFICITY: PITUITARY AND HYPOTHALAMUS.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@sdb-sdb.ch).
 CC -----
 DR EMBL: U60179; AAC50653.1; -.
 DR EMBL: U60181; AAC50654.1; -.
 DR GCRDB: GCR_1917; -.
 DR GCRDB: GCR_1918; -.
 DR MIM: 601898; -.
 DR InterPro: IPR0003905; GHS1_Receptor.
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR Pfam: PF00001; 7tm.1; 1.
 DR PRINTS: PR00237; GPCRHHODPSN.
 DR PROSITE: PS00237; G_PROTEIN_REC_F1_1; 1.
 DR PROSITE: PS50262; G_PROTEIN_REC_F1_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein;
 KM Alternative splicing.
 FT DOMAIN 1 40 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 41 66 1 (POTENTIAL).
 FT DOMAIN 67 72 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 73 96 2 (POTENTIAL).
 FT DOMAIN 97 117 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 118 139 3 (POTENTIAL).
 FT DOMAIN 140 162 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 163 183 4 (POTENTIAL).
 FT DOMAIN 184 211 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 212 235 5 (POTENTIAL).
 FT DOMAIN 236 263 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 264 285 6 (POTENTIAL).
 FT DOMAIN 286 302 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 303 326 7 (POTENTIAL).
 FT DOMAIN 327 366 CYTOPLASMIC (POTENTIAL).
 FT DISULFID 116 198 BY SIMILARITY.
 FT CARBOHYD 13 13 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 27 27 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPLIC 266 289 AVVAFATLCPMLPHVGRYLEKSK -> GGSORALRLSLAG
 FT PLSICLPLSL (IN ISOFORM 1B).
 FT MISSING (IN ISOFORM 1B).
 FT VARSPLIC 290 366
 FT SEQUENCE 366 AA; 41328 MW; D1B62710DA9DC06 CRC64;

Query Match 22.3%; Score 463.5; DB 1; Length 366;
 Best Local Similarity 34.9%; Pred. No. 3.4e-23;
 Matches 106; Conservative 66; Mismatches 107; Indels 25; Gaps 7;

DB 273 ILCLMPLPHVGRYLSKSFSEGLSLEIAISQYCNLVSFLYLSAAILPILYNIMSKRYV 332
 QY 328 AAFR 331
 DB 333 AVFR 336

RESULT 8
 ID NTR2_HUMAN STANDARD; PRT; 410 AA.
 AC 095665;
 DT 20-AUG-2001 (Rel. 40, Created)
 DT 20-AUG-2001 (Rel. 40, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE NEUTROTENSIN RECEPTOR TYPE 2 (NTR-2) (LEVOCABASTINE-SENSITIVE
 DE NEUTROTENSIN RECEPTOR) (NTR2 RECEPTOR).
 GN NTR2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=99066919; PubMed=9851594;
 RA Vita N., Oury-Donat F., Chalton P., Guillemot M., Kaghad M., Bachy A.,
 RA Thurneysen S., Garcia S., Polnot-Chazel C., Casellas P., Keane P.,
 RA Le Fur G., Maifrand J.P., Shoudry P., Caput D., Ferrara P.;
 Eur. J. Pharmacol. 360:265-272(1998).
 CC -1- FUNCTION: RECEPTOR FOR THE TRIDECAPEPTIDE NEUTROTENSIN. IT IS
 CC ASSOCIATED WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-
 CC CALCIUM SECOND MESSENGER SYSTEM.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC HIGHEST TO TACHYKININS RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@sdb-sdb.ch).
 CC -----
 DR EMBL: Y10148; CAA71233.1; -.
 DR MIM: 605538; -.
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR Pfam: PF00001; 7tm.1; 1.
 DR PRINTS: PR00237; GPCRHHODPSN.
 DR PROSITE: PS00237; G_PROTEIN_REC_F1_1; 1.
 DR PROSITE: PS50262; G_PROTEIN_REC_F1_2; 1.
 KW G-protein coupled receptor; Transmembrane; Lipoprotein; Palmitate.
 FT DOMAIN 1 32 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 33 55 1 (POTENTIAL).
 FT DOMAIN 56 64 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 65 87 2 (POTENTIAL).
 FT DOMAIN 88 109 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 110 131 3 (POTENTIAL).
 FT DOMAIN 132 154 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 155 176 4 (POTENTIAL).
 FT DOMAIN 177 217 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 218 237 5 (POTENTIAL).
 FT DOMAIN 238 297 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 298 318 6 (POTENTIAL).
 FT DOMAIN 319 337 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 338 358 7 (POTENTIAL).
 FT DOMAIN 359 410 CYTOPLASMIC (POTENTIAL).
 FT DISULFID 108 194 BY SIMILARITY.
 FT LIPID 377 PALMITATE (POTENTIAL).
 FT SEQUENCE 410 AA; 45413 MW; 8C3ADA22B615FD66 CRC64;


```

Query Match 21.1% Score 437 DB 1: Length 410:
Best Local Similarity 30.2% Pred. No. 1,9e-21:
Matches 108: Conservative 73: Mismatches 123: Indels 54: Gaps 11:

QY 44 YALIFLVGMNGLNLCVIVYHNOFLKT-PTNYVLFSLAVSDLLVLLCMPEIYE-MWNH 101
   |||:|:| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 39 YALIMALCAAMNALSIVNHLVKARAGRGRLRHHVYLLAGLILLVGLVPELVYEFWVFH 98
QY 102 YKPLFGPVGC---YKPTLFEYVCFASLSTPTVSVVEVYVAVLVHFRKLESTRRALRI 158
   |||:|:| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 99 YPMVFGDGLGCKGYF---VHELCAATVLSAAGLSAEKCLAVCOPLKARSLLTPRKTRWL 155
QY 159 LSLWVSFVSVDLSNPTSINGIKFQ-HPFNGSSVPSACATYTKPKMYNLLIQATSFLEY 217
   |||:|:| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 156 VALSMAASLGLALPMAVIMGOKHLELTADGEREPRASRYCTLVSTKATDVLFIQVNVLSF 215
QY 218 ILPMTLSLVL-----YILMG-----LRLKRDSESL-----EA 243
   |||:|:| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 216 VLPPLATLFLGLGVTVSHDLASCVPRSTPSTGSSSPRSRLLELSEGLLFTYMKKTFIIG 275
QY 244 NKVAVNIRPSRK-----SYTKMLEVLVLVAICVTPRHVDLRFESFY--EEMTESLAAY 296
   |||:|:| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 276 GCVSLVRRKIDVRRIRSLDORSYQVLRATVYMYICLPLRNARRLAKCYRDMDMTDPLNF 335
QY 297 ENLIHVSVEVFYSSAVNPITIIYLLNSRRFRAAFRNVSPTCKWCHPRNR----PQR 350
   |||:|:| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 336 YHFMVMTNLTFLVSSAVTPTPLLYNVAVSSSFRLKFLVAVSSLCGHNHPMKRLRPKQSP 393

RESULT 9
NTR2_RAT NTR2_RAT STANDARD; PRT; 416 AA.
AC 06384;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE NEUROTENSIN RECEPTOR TYPE 2 (NT-R-2) (HIGH-AFFINITY LEVOCABASTINE-SENSITIVE NEUROTENSIN RECEPTOR).
GN NTSR2 OR NTR2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Hypothalamus;
RX MEDLINE=96228041; PubMed=8647296;
RA Chalon P., Vila N., Kaghad M., Guillemont M., Bonin J.,
RA Delpech B., Le Fur G., Ferrara P., Caput D.;
RT "Molecular cloning of a levocabastine-sensitive neurotensin binding site.";
RL FEBS Lett. 386:91-94(1996).
CC -1- FUNCTION: RECEPTOR FOR THE TRIDECAPEPTIDE NEUROTENSIN. IT IS ASSOCIATED WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: ABUNDANT IN CORTEX AND HYPOTHALAMUS, AND LOWER LEVELS SEEN IN THE HEART AND INTESTINE.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED MAXIMALLY IN 7-DAY-OLD BRAIN AND EXPRESSION DECREASES PROGRESSIVELY UNTIL ADULTHOOD (35-DAY-OLD BRAIN).
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS. HIGHEST TO TACHIKININS RECEPTORS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sdb.ch/announce/ or send an email to license@isb-sdb.ch).

```

CC STRAIN-BALB/C; TISSUE-Brain;
 RX MEDLINE-96388216; PubMed-8795617;
 RA Mazella J., Boto J.-M., Guillemane E., Coppola T., Sarret P.,
 RT Vincent J.-P.;
 RT "Structure, functional expression, and cerebral localization of the
 RT levoocadherine-sensitive neurotensin/neuromedin N receptor from mouse
 RT brain";
 RL J. Neurosci. 16:5613-5620(1996).
 CC -1- FUNCTION: RECEPTOR FOR THE TRIDECAPEPTIDE NEURENSEN. IT IS
 CC ASSOCIATED WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-
 CC CALCIUM SECOND MESSENGER SYSTEM.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- TISSUE SPECIFICITY: EXPRESSED MAXIMALLY IN THE CEREBELLUM,
 CC HYPOTHALAMUS, PIRIFORM CORTEX AND NEOCORTIX OF ADULT BRAIN.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED POORLY IN 7-DAY-OLD BRAIN.
 CC EXPRESSION INCREASES AT DAY 15 TO REACH A MAXIMAL LEVEL IN 35-DAY-
 CC OLD BRAIN.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC HIGHEST TO TACHYKININS RECEPTORS.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U51908; AAB17285.1; -
 DR MGD: MGI:108018; Nstr2.
 DR InterPro: IPR000276; GPCR_Rhodopsin.
 DR Pfam: PF00001; 7tm_1; 2.
 DR PROSITE: PS00237; G-PROTEIN-RECEP.F1.1; 1.
 DR PROSITE: PS00262; G-PROTEIN-RECEP.F1.2; 1.
 KW G-protein coupled receptor; Transmembrane; Lipoprotein; Palmitate.
 FT DOMAIN 1 32
 FT TRANSMEM 33 55
 FT DOMAIN 56 64
 FT TRANSMEM 65 87
 FT DOMAIN 88 109
 FT TRANSMEM 110 131
 FT DOMAIN 132 154
 FT TRANSMEM 155 176
 FT DOMAIN 177 217
 FT TRANSMEM 218 238
 FT DOMAIN 239 298
 FT TRANSMEM 299 319
 FT DOMAIN 320 338
 FT TRANSMEM 339 359
 FT DOMAIN 360 417
 FT DISULFID 108 194
 FT LIPID 378
 SQ SEQUENCE 417 AA; 46537 MW; EBFDDDB650722DD CRC64;

Query Match 20.5%; Score 426.5; DB 1; Length 417;
 Best Local Similarity 31.0%; Pred. No. 8.8e-21;
 Matches 107; Conservative 64; Mismatches 119; Indels 55; Gaps 10;

QY 44 YALIFVGVGNLVCVVRHOTLKT-PTNYVLFSLAVSDLVLLGMLPEYVE-MWNN 101
 Db 39 YSLIFLGRAGNALSVHVKARTGRRLRYHVLALSALLLLLSVMEVYNFWSH 98
 QY 102 YPLFPGVGC---YFKTALFEVCAFSILSVTVSEVERVAIVHFFKLESTRRLARI 158
 Db 99 YPVVFGDLGGRGYF---VRELCAVATVLSVASLSAERCAVCOPLARARLLPRLRCRL 155
 QY 159 LSLVWSESVVFSLPNTSISIKFQ-HPNGSSVPGSATCTV----- 198
 Db 156 LSLVWVSLALPMVIMQKHEMERADDEPPASRVCTVLSRASRSSTFQVKRAGLL 215
 QY 199 TRMNVYNNILQATSF-----LFYILPMT-----LISVLYLMLRL 235

Db 216 RSPLMELTALINGITVNHVALYSQVPSASAOVNSIPSRLLESEGLGFTWRKTL 275
 QY 236 KRDESLFANVAVNIHPSKRSYTKMLFVLVFAICMPFHODLFESVEE--WTESL 293
 Db 276 GVOASLVRHKDASOIR--SLQHSQVLRALVAVYVICTWLPYHARRLWYCIIPDGWIDEL 333
 QY 294 AAVFNLHVYSGVFFYLSASVNPITVNLRSRFAFRNVPVSPYC 338
 Db 334 YDFHYHYVMTVNTLFYVSAVTPVLYNAVSSPFRKLFLLESLSLDC 378

RESULT 11
 TRFR_HUMAN
 ID TRFR_HUMAN STANDARD; PRT; 398 AA.
 AC P34981;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE THYROTROPIN-RELEASING HORMONE RECEPTOR (TRH-R) (THYROLIBERIN
 DE RECEPTOR).
 GN TRHR.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-93371401; PubMed-8395824;
 RA Matre V., Karlens H.E., Wright M.S., Lundell I., Fiedelheim K.,
 RA Gabrielsen O.S., Larhammer D., Gautvik K.M.;
 RT "Molecular cloning of a functional human thyrotropin-releasing
 RT hormone receptor";
 RL Biochem. Biophys. Res. Commun. 195:179-185(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Pituitary;
 RX MEDLINE-93384596; PubMed-8396925;
 RA Yamada M., Monden T., Satoh T., Satoh N., Murakami M., Iriuchijima T.,
 RA Kakegawa T., Mori M.;
 RT "Pituitary adenomas of patients with acromegaly express thyrotropin-
 RT releasing hormone receptor messenger RNA: cloning and functional
 RT expression of the human thyrotropin-releasing hormone receptor
 RT gene";
 RL Biochem. Biophys. Res. Commun. 195:737-745(1993).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-94063224; PubMed-8243797;
 RA Duthie S.M., Taylor P.L., Anderson L., Cook J., Eldne K.A.;
 RT "Cloning and functional characterisation of the human TRH receptor";
 RL Mol. Cell. Endocrinol. 95:R11-R15(1993).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-95002135; PubMed-7918619;
 RA Hinuma S., Hosoya M., Ogi K., Tanaka H., Nagai Y., Onda H.;
 RT "Molecular cloning and functional expression of a human thyrotropin-
 RT releasing hormone (TRH) receptor gene";
 RL Biochim. Biophys. Acta 1219:251-259(1994).
 RN [5]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-96355621; PubMed-8703031;
 RA Iwasaki T., Yamada M., Satoh T., Konaka S., Ren Y., Hashimoto K.,
 RA Kohga H., Kato Y., Mori M.;
 RT "Genomic organization and promoter function of the human thyrotropin-
 RT releasing hormone receptor gene";
 RL J. Biol. Chem. 271:22183-22188(1996).
 RN [6]
 RP SEQUENCE OF 1-263 FROM N.A.
 RC TISSUE-Placenta;
 RX MEDLINE-99101143; PubMed-9886052;
 RA Matre V., Hovring P.I., Orstavik S., Frengen E., Rian E.,
 RA Velickovic Z., Murray-McIntosh R.P., Gautvik K.M.;
 RT "Structural and functional organization of the gene encoding the
 RT human thyrotropin-releasing hormone receptor";

```

RL J. Neurochem. 72:40-50(1999).
CC -1- FUNCTION: RECEPTOR FOR THYROTROPIN-RELEASING HORMONE. THIS
CC RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE A
CC PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D16845; BAA04120.1; -.
DR EMBL: X75071; CA52965.1; -.
DR EMBL: X72089; CA50979.1; -.
DR EMBL: S75283; AAB32222.1; -.
DR EMBL: S75281; AAB32222.1; JOINED.
DR EMBL: D85376; BAA12796.1; -.
DR EMBL: D85375; BAA12796.1; JOINED.
DR EMBL: A1011701; CA09746.1; -.
DR PIR: JN0759; JN0759.
DR PIR: S40682; S40682.
DR PIR: JN0708; JN0708.
DR GCRDB: GCR_0765; -.
DR GCRDB: GCR_0799; -.
DR GCRDB: GCR_0880; -.
DR GCRDB: GCR_1111; -.
DR GCRDB: GCR_1148; -.
DR MIM: 188545; -.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PR00237; GPCR_Rhodopsn.
DR PRINTS: PR00751; THYROLIBRIN.
DR PROSITE: PS00337; G-PROTEIN_RECEP_F1_1; 1.
DR PROSITE: PS0262; G-PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
FT DOMAIN 1 28
FT TRANSMEM 29 51
FT DOMAIN 52 61
FT TRANSMEM 62 83
FT DOMAIN 84 99
FT TRANSMEM 100 121
FT DOMAIN 122 144
FT TRANSMEM 145 168
FT DOMAIN 169 193
FT TRANSMEM 194 215
FT DOMAIN 216 266
FT TRANSMEM 267 288
FT DOMAIN 289 296
FT TRANSMEM 297 319
FT DOMAIN 320 398
FT CARBOHD 3 3
FT CARBOHD 10 10
FT SEQUENCE 398 AA; 45084 MM; FE920B5FE293D3E CRC64;
SQ

```

Query Match 20.0%; Score 416; DB 1; Length 398;
 Best Local Similarity 31.5%; Pred. No. 4e-20;
 Matches 114; Conservative 66; Mismatches 120; Indels 62; Gaps 11;

```

DB 132 KA0FLCTFSRANKIIFVAFSTSLYCMWFLLDLINISTYKDAIY----SCGYKISRMV 187
OY 205 YNLIIGATSEFLYILPMTLSIVLYLMLGLR-----KRDESLANKAV 248
DB 188 YSPIYIMDGVYVPMILATVYLCIARILLENLPSPDKENSKTKWKNDSTHONTNLNV 247
OY 249 NIHR-----PSRKSVTNMFVLYLVFAICWTPFH---VDRLFSS--FVEEMTESLAIV 296
DB 248 MNSNRCFNSTVSRSKQVTMLAVVYLFLMLMPPRTLLVNVNSSLSPQENW----- 300
OY 297 ENLIHVSGVFYFLSSAVNPDIYNLSRFRFAFRNVNVSPTCKWCHPRHPOGPPAKII 356
DB 301 ---FLFCRCIYLNAINPVIYNLMSKFRFAFR-----KLCNCKOKPTEKPAHYSV 350
OY 357 FL 358
DB 351 AL 352

RESULT 12
TRFR_BOVIN STANDARD: PRT: 398 AA.
AC 046639:
DT 20-AUG-2001 (Rel. 40; Created)
DT 20-AUG-2001 (Rel. 40; Last sequence update)
DT 20-AUG-2001 (Rel. 40; Last annotation update)
DE THYROTROPIN-RELEASING HORMONE RECEPTOR (TRH-R) (THYROLIBERIN
  RECEPTOR).
GN TRHR.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HOLSTEIN;
RX MEDLINE=9815137; PubMed=9492373;
RA Takata M., Shimada Y., Ikeda A., Sekikawa K.;
RT "Molecular cloning of bovine thyrotropin-releasing hormone receptor
  gene.";
RL J. Vet. Med. Sci. 60:123-127(1998).
CC -1- FUNCTION: RECEPTOR FOR THYROTROPIN-RELEASING HORMONE. THIS
CC RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE A
CC PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D83964; BAA24069.1; -.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PR00751; THYROLIBRIN.
DR PROSITE: PS00337; G-PROTEIN_RECEP_F1_1; 1.
DR PROSITE: PS0262; G-PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
FT DOMAIN 1 28
FT TRANSMEM 29 51
FT DOMAIN 52 61
FT TRANSMEM 62 83
FT DOMAIN 84 99
FT TRANSMEM 100 121
FT DOMAIN 122 144
FT TRANSMEM 145 168
FT DOMAIN 169 193
FT TRANSMEM 194 215

```

FT DOMAIN 216 266 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 267 288 6 (POTENTIAL).
 FT DOMAIN 289 296 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 297 319 7 (POTENTIAL).
 FT DOMAIN 320 338 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 320 338 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 10 10 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 398 AA: 45165 MW: D9AF4B211A5701B8 CRC64:

Query Match 19.9%; Score 414; DB 1; Length 398;
 Best Local Similarity 31.1%; Pred. No. 5,3e-20;
 Matches 114; Conservative 69; Mismatches 115; Indels 68; Gaps 12;

QY 33 RSDLSPLVSAVAALIFLV---GVMGNLLVCVIVRHQTKPTNYLFSLAUSDLLVL 88
 DB 17 RAVVALEVOYVTLVLIIICGLGIVGINVYLVVMTKHKRTTNCYLSLAVADMLV 76
 QY 89 -LGMP--EIVEMHNPFLFPGVGCYFRTALFETVCPASILSVTVSVRYVAIVHP 144
 DB 77 AAGLPMTDSIYGSW-----VGVGCLCTITVLYLGINASSCSITAFTERIAICHP 131
 QY 145 RAKLESTRRALRIILWSFSVSESLPNTSINGIKFQHPNCGSSVPGSATCTVTYRPMV 204
 DB 132 KQOFLCTFSRAKKIIFVNAFTSYCMLWFLLDLNISTYKDAIV---SCGYKISRNY 187
 QY 205 YNLIQATSFILPMTLISVLYLM-----GLRLKROESLEANKVAVN 249
 DB 188 YSTIYIMDGVFTVPMILATVLYGFIARLFLNPIPSDEKNSNMKNSTHONK--N 244
 QY 250 IHR-----PSKRSYTKMLFVLVLFVFAICWTPFH--VDRLFFS--FVEEWTESL 293
 DB 245 LNSKTSNRYFNSTVSSRKQYTKMLAVVILFALMLMPRYTLVYVNSLSPFENM---- 300
 QY 294 AAVFNLIHVSGVFYLLSSAVNPITYNLSRRRAAFRNVSPTCKWCCHRRPQCPAQ 353
 DB 301 -----FLFECRICIYLSAINPIYVILMSQKRAFR-----KLCNCKOKRVEKPN 347
 QY 354 KIFLFT 359
 DB 348 YSAVLS 353

RESULT 13
 TRFR_CHICK STANDARD: PRT: 395 AA.
 ID TRFR_CHICK
 AC 093603;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE THYROTROPIN-RELEASING HORMONE RECEPTOR (TRH-R) (THYROLIBERIN RECEPTOR).
 DE RECEPTOR.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauromorpha; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Plutalry;
 RX MEDLINE=98344812; PubMed=9681487;
 RA Sun Y.M., Miller R.P., Ho H., Gershengorn M.C., Illing N.;
 RT Cloning and characterization of the chicken thyrotropin-releasing hormone receptor.;
 RL Endocrinology 139:3390-3398(1998).
 CC -1- FUNCTION: RECEPTOR FOR THYROTROPIN-RELEASING HORMONE. THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.1sb-sdb.ch/announce/> or send an email to license1sb-sdb.ch).

CC EMBL: Y18244; CAA77091.1; -
 DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR Pfam: PF00001; 7tm.1; 1.
 DR PRINTS: PR00237; GPCR_Rhodopsn.
 DR PRINTS: PR00751; THYROLIBERIN.
 DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; 1.
 DR PROSITE: PS00262; G_PROTEIN_RECEP_F1_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein.
 FT DOMAIN 1 30 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 31 53 1 (POTENTIAL).
 FT DOMAIN 54 63 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 64 85 2 (POTENTIAL).
 FT DOMAIN 86 101 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 102 123 3 (POTENTIAL).
 FT DOMAIN 124 146 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 147 170 4 (POTENTIAL).
 FT DOMAIN 171 195 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 196 217 5 (POTENTIAL).
 FT DOMAIN 218 268 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 269 290 6 (POTENTIAL).
 FT DOMAIN 291 298 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 299 321 7 (POTENTIAL).
 FT DOMAIN 322 395 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 3 3 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 10 10 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 395 AA: 44697 MW: 291D9BD2718723CB CRC64:

Query Match 19.9%; Score 413.5; DB 1; Length 395;
 Best Local Similarity 30.7%; Pred. No. 5,7e-20;
 Matches 119; Conservative 70; Mismatches 125; Indels 73; Gaps 14;

QY 42 VAAALIFLV---GVMGNLLVCVIVRHQTKPTNYLFSLAUSDLLVL-LGMP--LE 94
 DB 29 VTLVLIIICGLGIVGINVYLVVMTKHKRTTNCYLSLAVADMLVLAAGLPMTES 88
 QY 95 IYEMHNPFLFPGVGCYFRTALFETVCPASILSVTVSVRYVAIVHPRAKLESTRR 154
 DB 89 LYKSW-----VGVGCLCTITVLYLGINASSCSITAFTERIAICHPKAQFLCTFSR 143
 QY 155 ALRILSLWSFSVSESLPNTSINGIKFQHPNCGSSVPGSATCTVTYRPMYNNLIQATSR 214
 DB 144 AKKIIIFVWSFASVYCMLEFLLDLNIAVYKDTTV---SCGYKRSYSPIYMDFG 199
 QY 215 LFYILPMTLISVLYLMGRL-----KRDESLANKV-----AVNIH 251
 DB 200 IFYILPMTLISVLYLMGRL-----KRDESLANKV-----AVNIH 251
 QY 252 RPSKRSYTKMLFVLVLFVFAICWTPFH--VDRLFFS--FVEEWTESLAIVNLIHVSGV 306
 DB 260 IASRQVTKMLAVVILFALMLMPRYTLVYVNSLSPFENM-----FLFPCR 309
 QY 307 EYLLSAVNPITYNLSRRRAAFRNVSPTCKWCCHRRPQCPAQK1-----IFLTC 361
 DB 310 CIYLSAINPIYVILMSQKRAFRKCN-----CHLK-RDKKAPYSAVALNRYNKESD 363
 QY 362 HVLVELTDAGPQFGSSIHNTNLTRA 388
 DB 364 HFSSEIEDI-----TVNTYLSA 382

RESULT 14
 TRFR_MOUSE STANDARD: PRT: 393 AA.
 ID TRFR_MOUSE
 AC P21761;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-MAY-1991 (Rel. 18, Last sequence update)

DT 01-NOV-1997 (Rel. 35, last annotation update)
 DE THYROTROPIN-RELEASING HORMONE RECEPTOR (TRH-R) (THYROLIBERIN
 RECEPTOR).
 GN TRHR.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CC NCBI_TaxID:10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Pituitary;
 RX MEDLINE=91088548; PubMed=2175902;
 RA Straub R.E., Frech G.C., Joho R.H., Gershengorn M.C.;
 RT "Expression cloning of a cDNA encoding the mouse pituitary
 thyrotropin-releasing hormone receptor.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:9514-9518(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92381047; PubMed=1324930;
 RA Narayanan C.S., Fujimoto J., Geras-Raaka E., Gershengorn M.C.;
 RT "Regulation by thyrotropin-releasing hormone (TRH) of TRH receptor
 mRNA degradation in rat pituitary GH3 cells.";
 RL J. Biol. Chem. 267:17296-17303(1992).
 RN [3]
 RP SEQUENCE OF 332-393 FROM N.A.
 RC TISSUE-Pituitary;
 RX MEDLINE=97013702; PubMed=9156522;
 RA Jones K.E., Brubaker J.H., Chin W.W.;
 RT "An alternative splice variant of the mouse TRH receptor mRNA is the
 major form expressed in the mouse pituitary gland.";
 RL J. Mol. Endocrinol. 16:197-204(1996).
 CC -1- FUNCTION: RECEPTOR FOR THYROTROPIN-RELEASING HORMONE. THIS
 RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE A
 PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: M59611; AAA40480.1; -;
 DR EMBL: M94384; AAA40437.1; -;
 DR EMBL: L48936; AAA81559.1; -;
 DR PIR: A39251; A39251.
 DR GCRDB: GCR_0099; -;
 DR GCRDB: GCR_1613; -;
 DR MGD: MGI:98824; Trhr.
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR Pfam: PF00001; 7tm_1; 1.
 DR PRINTS: PRO0751; THYROLIBERIN.
 DR PROSITE: PS00237; G-PROTEIN_RECPT_FL_1; 1.
 DR PROSITE: PS50262; G-PROTEIN_RECPT_FL_2; 1.
 DR G-protein coupled receptor; Transmembrane; Glycoprotein.
 KW DOMAIN 1
 FT TRANSSEM 29 51
 FT TRANSSEM 52 61
 FT TRANSSEM 62 83
 FT TRANSSEM 84 99
 FT TRANSSEM 100 121
 FT TRANSSEM 122 144
 FT TRANSSEM 145 168
 FT TRANSSEM 169 193
 FT TRANSSEM 194 215
 FT TRANSSEM 216 266
 FT TRANSSEM 267 288
 FT TRANSSEM 289 296
 FT TRANSSEM 297 319
 FT DOMAIN 320 393
 CC CYTOPLASMIC (POTENTIAL).

FT CARBOHYD 3 3 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 10 10 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 393 AA; 44559 MM; 873987501A0CCEC CRC64;
 Query Match 19.8%; Score 412; DB 1; Length 393;
 Best Local Similarity 29.8%; Pred. No. 7,1e-20;
 Matches 114; Conservative 74; Mismatches 114; Indels 80; Gaps 13;
 QY 40 VSAVALIFL-----VGMGNLLVCMVIVRHOTKTPNITYLPSLAVSDLVLL-LG 90
 DB 20 VALEQVVTILLVIIICIGLGINIMVLVWRTKMRPTNICYLVSLAVDLMLVAVAG 79
 QY 91 MPL---EYEMKHNPFILPGVGCYEKTALEFVGCASLSTVTSVRYAVIAPPAK 147
 DB 80 LPINTDSIYGSW---YGVGVCCTIYLOYLGIMASCSSTAFETIATLCHPIAQ 134
 QY 148 LESTRRRLRLILSVSFSVPSLEPNTSIGHKFOHPNGSSVPSATCTVTKPMVYNL 207
 DB 135 FLCTFSRAKKIIFVWATFSIYCMFLDLINISTYKNAVY---SCGKISRNYS 190
 QY 208 IIOATSPFLYILPMTLISVLYLM-----GLRLKDESLKANKAVNINHR 252
 DB 191 IYLMDFGVYVPMIATLVYGFIRILFLNPISDPKENSIMKRNDSIHQNK-NLNPNA 249
 QY 253 -----PSKRSVTKMLFVLVFAICMPFH---VDRLFEFS--FVEEMTESLAVERN 298
 DB 250 TNRCNSVTSVSSKQYTKMLAVVILFALLMMPYRFLVAVNSFLSSPFGDN----- 300
 QY 299 LIHVSGVFYSSAVNPDIYMLSRFAAFRNVSPCKWCHPRHPOGPAQKII-- 356
 DB 301 -FLFRCRICIYNSAINVIYVILMSOKFRANR-----KICNCKQKPTKANAAYVAL 352
 QY 357 -----FTECHLVELTE 368
 DB 353 NYSVIESDPFSTEDLITVTD 374
 RESULT 15
 TRFR_SHEEP STANDARD: PRT; 398 AA.
 AC Q28596;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, last annotation update)
 DT 15-JUL-1998 (Rel. 36, last annotation update)
 DE THYROTROPIN-RELEASING HORMONE RECEPTOR (TRH-R) (THYROLIBERIN
 RECEPTOR).
 GN TRHR.
 OS Ovis aries (Sheep).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 CC Bovidae; Caprinae; Ovis.
 CC NCBI_TaxID=9940;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Pituitary;
 RX MEDLINE=97200775; PubMed=9048604;
 RA Bockmann J., Boeckers T.M., Winter C., Wiltkowski W., Winterhoff H.,
 RA Deufel T., Kreutz M.R.;
 RT "Thyrotropin expression in hypophyseal pars tuberalis-specific cells
 is 3,5,3'-triiodothyronine, thyrotropin-releasing hormone, and pit-1
 independent.";
 RL Endocrinology 138:1019-1028(1997).
 CC -1- FUNCTION: RECEPTOR FOR THYROTROPIN-RELEASING HORMONE. THIS
 RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE A
 PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL: X95285; CMA64606.1; -
DR GCRDB; GCR_1294; -
DR InterPro: IP000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PR00751; THYROLIBRIN.
DR PROSITE: PS00237; G_PROTEIN_RECP_F1_2; 1.
DR PROSITE: PS00262; G_PROTEIN_RECP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
FT DOMAIN 1 28 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 29 51 1 (POTENTIAL).
FT DOMAIN 52 61 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 62 83 2 (POTENTIAL).
FT DOMAIN 84 99 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 100 121 3 (POTENTIAL).
FT DOMAIN 122 144 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 145 168 4 (POTENTIAL).
FT DOMAIN 169 193 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 194 215 5 (POTENTIAL).
FT DOMAIN 216 266 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 267 288 6 (POTENTIAL).
FT DOMAIN 289 296 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 297 319 7 (POTENTIAL).
FT DOMAIN 320 398 CYTOPLASMIC (POTENTIAL).
FT CARBOHD 3 3 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 10 10 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 398 AA: 45088 MW: 375A311D3DD2A61A CRC64;

Query Match 19.8%; Score 412; DB 1; Length 398;
Best Local Similarity 31.5%; Pred No 7.2e-20;
Matches 115; Conservative 67; Mismatches 115; Indels 68; Gaps 12;

OY 33 RSDSLPVSVAVALFLV---GVMGNLVCMVIVRQTLKPTNYLFLSLAVSDLLVL 88
DB 17 RAVVALEYQVVTLLVLLVGLIGIVNIVLVVMTKMRPTNCYLVSLAVADLMVLV 76
OY 89 -LGMLP---EYEMHNINYPFLFGPGCYFKTALFETVCASILSYTVSVERVALVHP 144
DB 77 AAGLPRTIDSIYGSW---VYGVGCLCTITLYLQYLGINASSCSITAFIERIYAICHPI 131
OY 145 RAKLESTRRLRLIILSLVMSFSVFSLEPNTSINGIKFQHFPGSSVPGSATCTVKPMV 204
DB 132 KQOICTFSRAKKIIFVAFISYICMLWFLDLNISTYKDAIV---SCGYKISRNY 187
OY 205 VNLITQATSFLFYILPMLISLVLYLMGLRL-----KRDESLANKRAVAN 249
DB 188 YSPIYLMDFGVYVPMILATVLYGFIARILFLSPIDPKENSNTWKNDSTHONK---N 244
OY 250 IHR-----PSRKSVTMLFLVLYVEAICWTPFH---VDRLEFS--FVEEWTESL 293
DB 245 LMSKTSNRFTNSRKOVTMLAVVTLFALLMMPYRFLVAVNSFLSPQENW--- 300
OY 294 AAVFNLIHVSGVFYLLSSAVNPDIYNLISRRRAAFRNVAVSPTCKWCHPRHPOGPPAO 353
DB 301 -----FLLFRCICILYNSAINPVIYNLMSQKRAFR-----KLCNCKOKPYEKPRAN 347
OY 354 KTIIFL 358
DB 348 YSVAL 352

Search completed: April 22, 2002, 14:19:42
Job time: 145 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2002, 14:18:12 : Search time 21.56 Seconds
(without alignments)
412.282 Million cell updates/sec

Title: US-09-609-146-25
Perfect score: 2076
Sequence: 1 MGRLENSWIDHPLMKYILNS.....GQSSIHNTLTAPCAGEVP 395

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/prodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/prodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/prodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/prodata/2/1aa/PCJTUS.COMB.pep:*
6: /cgn2_6/prodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	465.5	22.4	364	4	US-09-077-675A-16	Sequence 16, Appl
2	465	22.4	353	4	US-09-077-675A-3	Sequence 3, Appl
3	463.5	22.3	361	4	US-09-077-675A-8	Sequence 8, Appl
4	463.5	22.3	366	4	US-09-077-675A-13	Sequence 13, Appl
5	437	21.1	410	3	US-08-858-876A-2	Sequence 2, Appl
6	437	21.1	410	4	US-09-472-880-2	Sequence 2, Appl
7	433.5	20.9	353	1	US-08-118-270-45	Sequence 45, Appl
8	433.5	20.9	353	5	PCT-US93-08528-45	Sequence 45, Appl
9	430.5	20.7	416	3	US-08-858-876A-4	Sequence 4, Appl
10	430.5	20.7	416	4	US-09-472-880-4	Sequence 4, Appl
11	423	20.4	302	4	US-09-077-675A-2	Sequence 2, Appl
12	421.5	20.3	302	4	US-09-077-675A-7	Sequence 7, Appl
13	416	20.0	398	2	US-08-288-663A-1	Sequence 1, Appl
14	412	19.8	393	1	US-07-629-1041-3	Sequence 3, Appl
15	400.5	19.3	400	3	US-08-889-108-8	Sequence 8, Appl
16	400.5	19.3	400	5	PCT-US94-10358-8	Sequence 8, Appl
17	399.5	19.2	384	2	US-08-103-170-10	Sequence 10, Appl
18	395.5	19.1	400	4	US-08-188-275A-2	Sequence 2, Appl
19	392	18.9	341	1	US-08-118-270-48	Sequence 48, Appl
20	392	18.9	341	5	PCT-US93-08528-48	Sequence 48, Appl
21	389.5	18.8	319	4	US-08-832-399-2	Sequence 2, Appl
22	389.5	18.8	319	4	US-09-372-498-2	Sequence 2, Appl
23	388.5	18.7	398	2	US-08-288-663A-15	Sequence 15, Appl
24	388.5	18.7	411	1	US-07-937-609-21	Sequence 21, Appl
25	380.5	18.3	411	4	US-08-029-170-21	Sequence 3, Appl
26	380	18.3	391	2	US-08-454-549-3	Sequence 3, Appl
27	380	18.3	391	3	US-08-454-552-3	Sequence 3, Appl

28	380	18.3	398	1	US-08-149-093A-5	Sequence 5, Appl
29	380	18.3	398	2	US-08-911-245-5	Sequence 5, Appl
30	380	18.3	398	2	US-08-889-108-2	Sequence 2, Appl
31	380	18.3	398	4	US-08-120-601B-2	Sequence 2, Appl
32	380	18.3	398	4	US-08-188-275A-3	Sequence 3, Appl
33	380	18.3	398	4	US-08-387-707-16	Sequence 16, Appl
34	380	18.3	398	4	US-09-510-473-5	Sequence 5, Appl
35	380	18.3	398	5	PCT-US94-10358-2	Sequence 2, Appl
36	374	18.0	391	3	US-08-676-351-4	Sequence 4, Appl
37	374	18.0	398	3	US-09-170-331-5	Sequence 5, Appl
38	373	18.0	387	1	US-08-196-989B-14	Sequence 14, Appl
39	373	18.0	387	2	US-08-760-936-14	Sequence 10, Appl
40	371	17.9	451	4	US-08-430-286A-10	Sequence 10, Appl
41	370	17.8	356	4	US-08-430-286A-2	Sequence 2, Appl
42	368	17.7	369	1	US-07-937-609-19	Sequence 19, Appl
43	368	17.7	369	4	US-08-029-170-19	Sequence 19, Appl
44	366.5	17.7	465	4	US-08-090-369-1	Sequence 1, Appl
45	363.5	17.5	384	3	US-09-071-434-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-09-077-675A-16
Sequence 16, Application US/09077675A
Patent No. 6242199
GENERAL INFORMATION:
APPLICANT: Pai, Lee-Yuh
APPLICANT: Feighner, Scott C.
APPLICANT: Howard, Andrew D.
APPLICANT: Pong, Sheng-Shung
APPLICANT: Van Der Ploeg, Leonardus H.T.
TITLE OF INVENTION: RECEPTOR ASSAY
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000, 126 E. Lincoln Ave.
CITY: Rahway
STATE: NJ
COUNTRY: USA
ZIP: 07065-0900
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/077,675A
FILING DATE: 3-JUN-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Cocuzzo, Anna L.
REGISTRATION NUMBER: 42,452
REFERENCE/DOCKET NUMBER: 19590P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 732-594-1273
TELEFAX: 732-594-4720
TELEX:
INFORMATION FOR SEQ. ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 364 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-077-675A-16
Query Match 22.4%; Score 465.5; DB 4; Length 364;

Best Local Similarity 34.6%; Pred. No. 5.2e-31;
Matches 104; Conservative 65; Mismatches 107; Indels 25; Gaps 6;

QY 40 VSAVYALIFLVGVGNLJYCMVIVRHOTLKPTNYVLFSLAVSDLLVLLGMPLEIEM 99
Db 45 VTATCVALLFVVGISGNLTLVSVRFELRTTNLYLSMAFSDLLFLC-MPLDLRLW 103
QY 100 HNYFELGPGVCEYKTLAFETVCASILSVTVSVERYVAIVHFRKLESTRRARIL 159
Db 104 QYRWMNGLNCKLFQGVSESCYATVLTITALSVERYPALCFPLRAKVVYTKRVLVI 163
QY 160 SLWSESVSVSLPMTSHGKIFQHFPGSSVPGSATCTVTK-----PMVYLLIT 209
Db 164 LVIAVAFCSGPIFLVAVGEH---NGTDPDRTNECRATEFAVRSGLLTVWVW----- 215
QY 210 QATSEFLYILPMTLISVLYLMGLRLKRDSELEANKVAVINHRPSRKYTKMLFVLVLF 269
Db 216 ---SVFFELPVFCITVLYSLIGRKLRRRGDA--VGASLRDQNHQYVKMLAVVFAF 270
QY 270 AICWTPHYDRLPFS-EVEWTESTLAAVFNLIVHVGCVFFYLSAVNPITVNLRRRFA 328
Db 271 ILCWLPYHGVGRYLSKSPGSELEIAQISOYCNLVSFVLYLSAANPILYNIMSKRYV 330
QY 329 A 329
Db 331 A 331

RESULT 2
US-09-077-675A-3
Sequence 3, Application US/09077675A
Patent No. 6242199

GENERAL INFORMATION:

APPLICANT: Pal, Lee-Yuh
APPLICANT: Feighner, Scott C.
APPLICANT: Howard, Andrew D.
APPLICANT: Pong, Sheng-Shung
APPLICANT: Van Der Ploeg, Leonardus H.T.
TITLE OF INVENTION: RECEPTOR ASSAY
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000, 126 E. Lincoln Ave.
CITY: Rahway
STATE: NJ
COUNTRY: USA
ZIP: 07065-0900

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/077, 675A
FILING DATE: 3-JUN-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: Cocuzzo, Anna L.
REGISTRATION NUMBER: 42,452
REFERENCE/DOCKET NUMBER: 19590P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 732-594-1273
TELEFAX: 732-594-4720

TELEX:

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:
LENGTH: 353 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: protein
US-09-077-675A-3

Query Match 22.4%; Score 465; DB 4; Length 353;
Best Local Similarity 34.2%; Pred. No. 5.5e-31;
Matches 103; Conservative 60; Mismatches 106; Indels 24; Gaps 6;

QY 40 VSAVYALIFLVGVGNLJYCMVIVRHOTLKPTNYVLFSLAVSDLLVLLGMPLEIEM 99
Db 33 VTATCVALLFVVGISGNLTLVSVRFELRTTNLYLSMAFSDLLFLC-MPLDLRLW 91
QY 100 HNYFELGPGVCEYKTLAFETVCASILSVTVSVERYVAIVHFRKLESTRRARIL 159
Db 92 QYRWMNGLNCKLFQGVSESCYATVLTITALSVERYPALCFPLRAKVVYTKRVLVI 151
QY 160 SLWSESVSVSLPMTSHGKIFQHFPGSSVPGSATCTVTK-----PMVYLLIT 209
Db 152 LVIAVAFCSGPIFLVAVGEH---NGTDPDRTNECRATEFAVRSGLLTVWVW----- 203
QY 210 QATSEFLYILPMTLISVLYLMGLRLKRDSELEANKVAVINHRPSRKYTKMLFVLVLF 269
Db 204 ---SVFFELPVFCITVLYSLIGRKLRRRGDA--AVGSSLRDQNHQYVKMLAVVFAF 259
QY 270 AICWTPHYDRLPFS-EVEWTESTLAAVFNLIVHVGCVFFYLSAVNPITVNLRRRFA 328
Db 260 ILCWLPYHGVGRYLSKSPGSELEIAQISOYCNLVSFVLYLSAANPILYNIMSKRYV 319
QY 329 A 329
Db 320 A 320

RESULT 3
US-09-077-675A-8
Sequence 8, Application US/09077675A
Patent No. 6242199

GENERAL INFORMATION:

APPLICANT: Pal, Lee-Yuh
APPLICANT: Feighner, Scott C.
APPLICANT: Howard, Andrew D.
APPLICANT: Pong, Sheng-Shung
APPLICANT: Van Der Ploeg, Leonardus H.T.
TITLE OF INVENTION: RECEPTOR ASSAY
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000, 126 E. Lincoln Ave.
CITY: Rahway
STATE: NJ
COUNTRY: USA
ZIP: 07065-0900

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/077, 675A
FILING DATE: 3-JUN-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: Cocuzzo, Anna L.
REGISTRATION NUMBER: 42,452
REFERENCE/DOCKET NUMBER: 19590P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 732-594-1273
TELEFAX: 732-594-4720

TELEX:

INFORMATION FOR SEQ ID NO: 8:

ATTORNEY/AGENT INFORMATION:
NAME: Player, William E.
REGISTRATION NUMBER: 31,049
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 410 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-858-876A-2

Query Match 21.1%; Score 437; DB 3; Length 410;
Best Local Similarity 30.2%; Pred. No. 1.3e-28;
Matches 108; Conservative 73; Mismatches 123; Indels 54; Gaps 11;

QY 44 YALIFLVGNGNLVCMVYVRHOTLKT-PTNYLFSLANSDDLVLGLMPLEIYE-MMHN 101
DB 39 YALIMALGAAGNALSVHVVYLKARAGRGLRHHVLSLALAGLLLVGVPELVSEVWFH 98
QY 102 YPFLFGPVC---YFKTALFETVCFAISLVTVSVERVVAIVHPRAKLESTRRALRT 158
DB 99 YPWFEDLCGRGYF---VHELCAVATVLSVAGLSAECLAVCQPLRAKSLTPRRTRWL 155
QY 159 LSLWSESVFSLPNTSINGIKFQ-HFPNGSSVGSATCTVTKPMWYNLLIOATSEFLFY 217
DB 156 VALSMASIGLALPMVINGOKHELETADGEPPASRVCTVLSRTALQYFIOVNVLSF 215
QY 218 ILPMTLISLV-----YLMG-----LRKDESL-----EA 243
DB 216 VLPLATLAFNGVTVSHLALCSQVSTSPGSSPFSRLLELSEGLISPIVAKKTFIOG 275
QY 244 NKAVNINHRPSRK-----SVTKMLFVLVLEAFICWTPFHVDRLFFSFV---EEMTESLA 296
DB 276 GOVSLVRHKDVRIRSLQNSVOYLRAIVMYTCMLPYHARRMLCYTPDDAWTIDPLYNF 335
QY 297 FNLIHVSGVFYLSAANPDIYNLSRRFRAFRNVSPCTCKMCHPRNR----POGP 350
DB 336 YHFFYVNTNLFYVSSAVPRLYNAVSSSRKFLFLEAVSLGCEHHPMKRLPRKQSP 393

RESULT 6
US-09-472-880-2
Sequence 2, Application US/09472880
Patent No. 6274333
GENERAL INFORMATION:
APPLICANT: Daniel CAPUT
PASCAL CHALON
PASCAL FERRARA
VILTA NATALIO
TITLE OF INVENTION: Type 2 Neurotensin Receptor
(hmr-R2)
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Jacobson, Price, Holman & Stern, PLLC
STREET: 400 Seventh Street
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/472,880
FILING DATE: 28-Dec-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR 9723204
FILING DATE: 17-MAR-1997

ATTORNEY/AGENT INFORMATION:
NAME: Player, William E.
REGISTRATION NUMBER: 31,049
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 410 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-472-880-2

Query Match 21.1%; Score 437; DB 4; Length 410;
Best Local Similarity 30.2%; Pred. No. 1.3e-28;
Matches 108; Conservative 73; Mismatches 123; Indels 54; Gaps 11;

QY 44 YALIFLVGNGNLVCMVYVRHOTLKT-PTNYLFSLANSDDLVLGLMPLEIYE-MMHN 101
DB 39 YALIMALGAAGNALSVHVVYLKARAGRGLRHHVLSLALAGLLLVGVPELVSEVWFH 98
QY 102 YPFLFGPVC---YFKTALFETVCFAISLVTVSVERVVAIVHPRAKLESTRRALRT 158
DB 99 YPWFEDLCGRGYF---VHELCAVATVLSVAGLSAECLAVCQPLRAKSLTPRRTRWL 155
QY 159 LSLWSESVFSLPNTSINGIKFQ-HFPNGSSVGSATCTVTKPMWYNLLIOATSEFLFY 217
DB 156 VALSMASIGLALPMVINGOKHELETADGEPPASRVCTVLSRTALQYFIOVNVLSF 215
QY 218 ILPMTLISLV-----YLMG-----LRKDESL-----EA 243
DB 216 VLPLATLAFNGVTVSHLALCSQVSTSPGSSPFSRLLELSEGLISPIVAKKTFIOG 275
QY 244 NKAVNINHRPSRK-----SVTKMLFVLVLEAFICWTPFHVDRLFFSFV---EEMTESLA 296
DB 276 GOVSLVRHKDVRIRSLQNSVOYLRAIVMYTCMLPYHARRMLCYTPDDAWTIDPLYNF 335
QY 297 FNLIHVSGVFYLSAANPDIYNLSRRFRAFRNVSPCTCKMCHPRNR----POGP 350
DB 336 YHFFYVNTNLFYVSSAVPRLYNAVSSSRKFLFLEAVSLGCEHHPMKRLPRKQSP 393

RESULT 7
US-08-118-270-45
Sequence 45, Application US/08118270
Patent No. 5508364
GENERAL INFORMATION:
APPLICANT: Murphy, Randall B.
APPLICANT: Schuster, David I.
TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
NUMBER OF SEQUENCES: 348
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NETMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/118,270
FILING DATE: 09-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/943,236
FILING DATE: 10-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Townsend, Kevin G.
REGISTRATION NUMBER: 34,033

APPLICANT: HINUMA, Shuji
 APPLICANT: HOSOYA, Masaki
 APPLICANT: ONODA, Haruo
 TITLE OF INVENTION: HUMAN TRH RECEPTOR, ITS PRODUCTION
 TITLE OF INVENTION: AND USE
 NUMBER OF SEQUENCES: 15
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSMAN, LLP
 STREET: 130 Water Street
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02109
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FASTSO for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/288,663A
 FILING DATE: 09-AUG-1994
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:

```

: APPLICATION NUMBER: 198309/1993
: FILING DATE: 10-AUG-1993
: APPLICATION NUMBER: 286986/1993
: FILING DATE: 16-NOV-1993
: APPLICATION NUMBER: 325215/1993
: FILING DATE: 22-DEC-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: Resnick, David S
: REGISTRATION NUMBER: 34,235
: REFERENCE/DOCKET NUMBER: 44612
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 617-523-3400
: TELEFAX: 617-523-6440
:
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 398 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-288-663A-1

```

```

Query Match          20.0%; Score 416; DB 2; Length 398;
Best Local Similarity 31.5%; Pred. No. 6,8e-27;
Matches 114; Conservative 66; Mismatches 120; Indels 62; Gaps 11;

```

```

QY 33 RSDLSLPVSAVALIFLY-----GYMGNLVCMYIVRHQTLKTPPNYYLSLAVSDDLVL 88
DB 17 RAAVALEYQVVTLLVILICGLGIVGIMVAVLMRTKHKRTPTNCLVSLAVADLVAV 76
QY 89 -LCMPL--EYEMMHNPFLPCPGCYFKTALFETVCFASISLVTVSERYVAIVHPF 144
DB 77 AAGLPNTDSIGSW-----VGVGCLCTITLYOYGINASSCSITAFTERIATICHPI 131
QY 145 RALESTRRALRLSLVMSFSVFSLPNTSINGIKFOHPNGSSVPGSATCTVTRPMVY 204
DB 132 KAQFLCFESRAKKIIFVMAFTSLYCMLEFLLDINISTYKDAIVI-----SCGYKISRNY 187
QY 205 YNLIQATSFYILPMTLSVLYLMGLRL-----KRDESLANKVAV 248
DB 188 YSPYLMDFGVVPMILATVYGFARILFLNPIPSDPKSKMKNDSIQNTNLN 247
QY 249 NIHR-----PSRKSVTKMLVLYVFAICWTPFH---VDRLFFS--FVEEMTESLA 296
DB 248 NTRNCSTVSSRKQVTKMLAVVILFALLMMPYRLVAVNSLSPQENW----- 300
QY 297 ENLHVSGVFYLLSSAVNPDIYNLLSRFRRAFRNVVSPCKWCHPRHRPOGPPAKII 356
DB 301 -FLFECRICIYLNSAIPVIYNLMQKFRAPFR-----KLCNCKOKPTEKAPANYSV 350
QY 357 FL 358
DB 351 AL 352

```

```

RESULT 14
US-07-629-1041-3
: Sequence 3, Application US/076291041
: Patent No. 5288621
: GENERAL INFORMATION:
: APPLICANT: Gershengorn, Marvin C
: TITLE OF INVENTION: PITUITARY TRH RECEPTOR
: NUMBER OF SEQUENCES: 3
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Yahwak & Associates
: STREET: 25 Skytop Drive
: CITY: Trumbull
: STATE: Connecticut
: COUNTRY: USA
: ZIP: 06611
: COMPUTER READABLE FORM:

```

```

: MEDIUM TYPE: Floppy disk
: COMPUTER: Macintosh
: OPERATING SYSTEM: MS-DOS
: SOFTWARE: Microsoft Word 3.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/629,1041
: FILING DATE: 19901214
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: George M. Yahwak
: REGISTRATION NUMBER: 26,824
: REFERENCE/DOCKET NUMBER: CRF D - 995
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (203)268-1951
: TELEFAX: (203)268-1951
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 393 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: US-07-629-1041-3

```

```

Query Match          19.8%; Score 412; DB 1; Length 393;
Best Local Similarity 29.8%; Pred. No. 1,4e-26;
Matches 114; Conservative 74; Mismatches 114; Indels 80; Gaps 13;

```

```

QY 40 VSAVALIFL-----VGMGNLVCYIVRHQTLKTPPNYYLSLAVSDDLVL-LG 90
DB 20 VALEYQVVTLLVILICGLGIVGIMVAVLMRTKHKRTPTNCLVSLAVADLVAVAG 79
QY 91 MPL--EYEMMHNPFLPCPGCYFKTALFETVCFASISLVTVSERYVAIVHPRAK 147
DB 80 LPNTDSIGSW-----VGVGCLCTITLYOYGINASSCSITAFTERIATICHPIKAQ 134
QY 148 LESTRRRALRLSLVMSFSVFSLPNTSINGIKFOHPNGSSVPGSATCTVTRPMVYNL 207
DB 135 FLCFESRAKKIIFVMAFTSLYCMLEFLLDINISTYKKNVY-----SCGYKISRNYSP 190
QY 208 TIQATSFYILPMTLSVLYLM-----GLRLKRDSELANKVAVNIHR 252
DB 191 IYLMDFGVVPMILATVYGFARILFLNPIPSDPKSKMKNDSIQNK-NLNLNA 249
QY 253 -----PSRKSVTKMLVLYVFAICWTPFH---VDRLFFS--FVEEMTESLA 298
DB 250 TNRNCSTVSSRKQVTKMLAVVILFALLMMPYRLVAVNSLSPQENW----- 300
QY 299 ENLHVSGVFYLLSSAVNPDIYNLLSRFRRAFRNVVSPCKWCHPRHRPOGPPAKII 356
DB 301 -FLFECRICIYLNSAIPVIYNLMQKFRAPFR-----KLCNCKOKPTEKAPANYSV 352
QY 357 -----FLTECHLVELTE 368
DB 353 NYSVYKESDRSTELIEDITVTD 374

```

```

RESULT 15
US-08-889-108-8
: Sequence 8, Application US/08889108
: Patent No. 6103492
: GENERAL INFORMATION:
: APPLICANT: Yu, Lei
: TITLE OF INVENTION: Mu Opioid Receptors: Compositions and Methods
: NUMBER OF SEQUENCES: 17
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Arnold, White & Durkee
: STREET: P. O. Box 4433
: CITY: Houston
: STATE: TX
: COUNTRY: USA
: ZIP: 77210-4433
: COMPUTER READABLE FORM:

```

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/889,108
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/305,518
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Wilson, Mark B.
REGISTRATION NUMBER: 37,259
REFERENCE/DOCKET NUMBER: INDA005\WIM
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512-418-3000
TELEFAX: 512-474-7577
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 400 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US: 08-889-108-8

Query Match 19.3% Score 400.5; DB 3; Length 400;
Best Local Similarity 30.1%; Pred. No. 1.3e-25;

Matches 106; Conservative 66; Mismatches 133; Indels 47; Gaps 11;

QY 8 SWHDPMLKYLNSTEYLAHLGPKRSDL-----SLPVSVAVALIFLY 50
DB 30 SWVN-----LSHLEGNLSDPCGPNRTDLGGRDSLCPPTGSPSMITAITIMALYSIVCV 83
QY 51 GVMGNLVCMVYVRHOTLKTPTNYLFSLAVALDLVLLGMPLE-IYEMHNYPLFGPV 109
C 84 GLFGNFMVYIVRTKMTATNITYLALDALATST-LPQSVNYLMGTWP--FGTI 140
QY 110 GCYRTALFETVCFASISLVTTSVERVAVIHPFRALKESTRRRALRLISLVSFSVVF 169
DB 141 LKIVISIDVYVMFTSIFTLCTMSVDRIAVCHPVKALDFRTPRNAKLIINCNMILSSAI 200
QY 170 SLPTNSIGIKFOHFPNNSVPGSATCTVT--KPMVYNLIIOATSFLE-YLPMTLISV 226
DB 201 GLPWFMAATTYKRO-----GSIDCTLTFPSHPTWYENLKICVFIFAFIMPLIITV 252
QY 227 LYYLMLGLRKDESLANKVA-VNIHRSRKSVMKMLFVLVLFALICWTPZHVDRLFFSF 285
DB 253 CYGLMILRLKSVRLSGSKENDRLNR-----ITRMVLVVAVFIVCWTPLHIYIIRKAL 307
QY 286 VEEVTESLAVALNLIHVVSGVFYLLSSAVNPDIYNLRSRFRFAFRNVVSP 337
DB 308 VTIPETTFOTVSMHFCIALG---YTNGLNPLVLAFLDENKRCRCFRFCIPT 356

Search completed: April 22, 2002, 14:18:13
Job time: 326 sec

THIS PAGE BLANK (uspro)